RICE UNIVERSITY

# Mid-IR Spectral Investigation of Normal and Malignant Breast and Cervical Tissue Samples Using a Quantum Cascade Laser-Based Microscope

by

**Paul Haugen**

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

**Doctor of Philosophy**

APPROVED, THESIS COMMITTEE

Rebekah Drezek, PhD, Thesis Director
Professor of Bioengineering and Electrical and Computer Engineering

Frank Tittel, PhD, Committee Chair
J.S. Abercrombie Professor of Electrical and Computer Engineering

Junichiro Kono, PhD
Professor of Electrical and Computer Engineering

HOUSTON, TEXAS
January 2017

# Abstract


## Mid-IR Spectral Investigation of Normal and Malignant Breast and Cervical Tissue Samples Using a Quantum Cascade Laser-Based Microscope

by

**Paul Haugen**

Mid-infrared (MIR) spectroscopy has been a tool used to identify specific features of normal and malignant tissue samples by utilizing MIR characteristics, specifically in the "fingerprint" region. The fingerprint region is a biologically significant spectral region typically identified between 1500 and 500 $cm^{-1}$. MIR spectroscopy can be used to study molecular changes and variations occurring in samples, which can then be used to fingerprint specific spectral characteristics and biomarkers in order to categorize the specimens. The most common instruments currently used in this analysis are Fourier transform infrared (FTIR) spectrometers, although properties inherent in these instruments, such as slow data collection time and an inability to specify sample location for the spectral data collection, have placed a ceiling on the clinical practicality of their use for specimen classification and identification. In this thesis, we use a prototype of an infrared hyperspectral imaging microscopy platform based around tunable quantum cascade laser (QCL) technology that has a spectral coverage from 1800-900 $cm^{-1}$. The quantum cascade lasers are coupled with a series of MIR refractive objectives and an uncooled microbolometer camera. The speed of spectral imaging improves to 30 frames per

second, and the high magnification objective has a 1.34 μm pixel resolution with a 0.70 numerical aperture and 4.3 μm spatial resolution. We are able to specify data collection at specific discrete wavelengths as opposed to the full spectrum, which improves the data collection time and de-clutters the data for analysis expediency. Finally, we perform spectral imaging real-time, which aides in selecting precise regions of interest on the target sample. This thesis demonstrates the advantages of exploiting the capabilities of the QCL microscope to advance MIR spectroscopy in the identification of distinguishing traits of normal and malignant breast and cervical tissue samples.

# Acknowledgments

I would like to thank the financial support I have received. The National Science Foundation (NSF), the NSF-affiliated Mid-InfraRed Technologies for Health and the Environment (MIRTHE), and Rice University's Drezek Lab supported funding for this project.

I would like to thank my fellow lab-mates, the following of whom put up with my weird hours of operation and allowed me to bounce many ideas (not all good) off them: Vishu Asthana, Lissett Bickford, Allen Chen, Elizabeth Figueroa, Ying Hu, Laura Kennedy, Robert Langsner, Nastassja Lewinski, Adam Lin, Joao Paolo Mattos, Vengadesan Nammalvar, Sterling Noelck, Emily Reiser, and Joseph Young. I would also like to thank my thesis committee, Dr. Frank Tittel and Dr. Junichiro Kono, for their willingness to work with me through my master's work and assist with my dissertation.

I want to give special thanks to my advisor, Dr. Rebekah Drezek. I don't think she realizes the impact she has had in the direction of my life and my family's life. My first day on campus, she took the time to sit down with me and offer suggestions for both my master's work and PhD work; she also gave me a desk and computer and allowed me to do research in her lab. Not only was I not in her lab yet, she had never met or talked to me before. Her guidance and wisdom has been crucial in my academic life, which has then impacted both my personal and professional life and has opened up doors that were not possible before. Her patience and insight in dealing with my unique situations was unparalleled, and I don't think I can express

my gratitude enough to her. Thank you, Dr. Drezek, for all the above-and-beyond effort you put into working with me.

Finally, I would like to especially thank my wife Jill of 14 years and my three children, Hadley (9), Jonah (7), and Nissa (1). They have put up with my many long hours tucked away in the office at home or sitting at a coffee shop trying to complete my dissertation. My wife is the one that really encouraged me to continue with my education and completed my PhD. She spent many days and evenings covering for my slack while I was preoccupied, and my children were always trying so hard to help me with my thesis. (The best-written parts of my thesis are probably the sections they wrote!) I won't forget when my oldest daughter, Hadley, had her "packpack" on and told me that I was going to school to learn about numbers just like her, or when Jonah would make a computer out of paper and cardboard so he could sit by me and work, too, or when Nissa would sit on my lap and "help" me type. I never would have been able to complete my dissertation without the complete support and backing I had from Jill and from Hadley, Jonah, and Nissa.

# Contents

# List of Figures

# List of Tables

Chapter 1

# Introduction to Mid-IR Cancer Detection Using FTIR Spectroscopy

Cancer is the second leading cause of death in the United States, second only to heart disease. 8.5% of all Americans have been diagnosed with cancer. According to the Center for Disease Control and Prevention, cancer accounts for 23% of the deaths, or about 585,000 deaths annually. Heart disease is responsible for 24% of the deaths in the U.S., or about 610,000 deaths annually. The third leading cause of death, chronic lower respiratory disease, is accountable for less than 6% of the annual deaths, so there is a large discrepancy between the top two killers and the rest of the field. One alarming statistic of cancer is that the rate is continually increasing every year and is slowly closing the gap with heart disease. The survival rate for people diagnosed with cancer greatly increases with early detection.[1,2]

Fourier transform infrared (FTIR) spectroscopy is a method that shows promise in either replacing or assisting current cancer detection techniques by looking at absorption levels at various wavelengths and "fingerprinting" the cancer cells. A unique benefit of this method is that contrast agents are not needed; however, several factors affect the accuracy and repeatability of FTIR spectroscopy. Specifically, no current standards exist for the sample preparations, data acquisition processes (particularly with the attenuated total reflectance [ATR] attachment), or data processing. Thus, the data results from different research groups are difficult to compare. We initially surveyed this method of cancer detection and analysis to document and evaluate the current techniques being used. Our investigation was based on published papers and research from various groups working in this field. In addition, our own data samples and techniques were referenced and documented.

FTIR spectroscopy shows promise as a technique for early cancer detection. This is a technique to study molecular changes and variations occurring between cancerous and normal samples. Not only does it show potential for segregating between normal and cancer cells, but it also shows potential in differentiating between types of cancer, which can aid in cancer treatment.[3]

## 1.1. FTIR Spectroscopy

### 1.1.1. Sample Preparation

We maintained up to six cell lines at a time, which involved planting and growing the cells and changing the media several times per week. When the cells covered a significant amount of space in the flask in which they were growing, we then trypsinized the cells, which detach the cells from the bottom of the flask. The cells had to be trypsinized before they covered too much of the flask, however, or the cells would begin developing unhealthy attributes and eventually start dying.

After the cells were trypsinized, we spun the cells down to create a pellet of cells. We typically spun the cells down in a centrifuge at 125 g for 5 minutes; we were then able to remove most of the media and trypsin from the cell pellet. Special care had to be taken to properly rinse the cells in order to prevent foreign material (such as media or trypsin) from affecting the spectra during the data acquisition process. Initially, we used a phosphate buffered saline (PBS) solution to rinse the cells; however, it was soon shown that this method allowed the formation of crystals on the cells after the rinsing process. Thus, deionized water (DI water) proved to be a more effective rinsing method.

After the cells were properly rinsed with deionized water, the best method we found was to fix the cells, which essentially prevented the cells' properties from changing after they were dried on the slides. Finally, they were rinsed one more time in deionized water, and five drops of 10 μl were evenly deposited onto each

slide and allowed to air dry. We were typically using Mid-IR slides from Kevley Technologies.

We tried several other methods of sample preparation, such as cryofreezing or growing the cells on the actual slides, but these methods were not as effective as the technique described above.

In addition to the cell sample preparation method, other tasks we performed included preparing tissue samples by slicing frozen breast cancer and preparing slides for further evaluation. Overall good general practice skills were also developed for working in a bioengineering laboratory, including the cultivation of good practices for cell culture techniques.

### 1.1.2. Data Acquisition

We used a Nicolet Nexus 670 FTIR spectrometer with a liquid nitrogen-cooled mercury cadmium telluride (MCT) detector and a potassium bromide (KBr) beamsplitter to collect our sample spectra. We also incorporated an attenuated total reflectance (ATR) attachment for our data collection. See Figure 1.1 for sample spectra we collected for four of our cell lines.

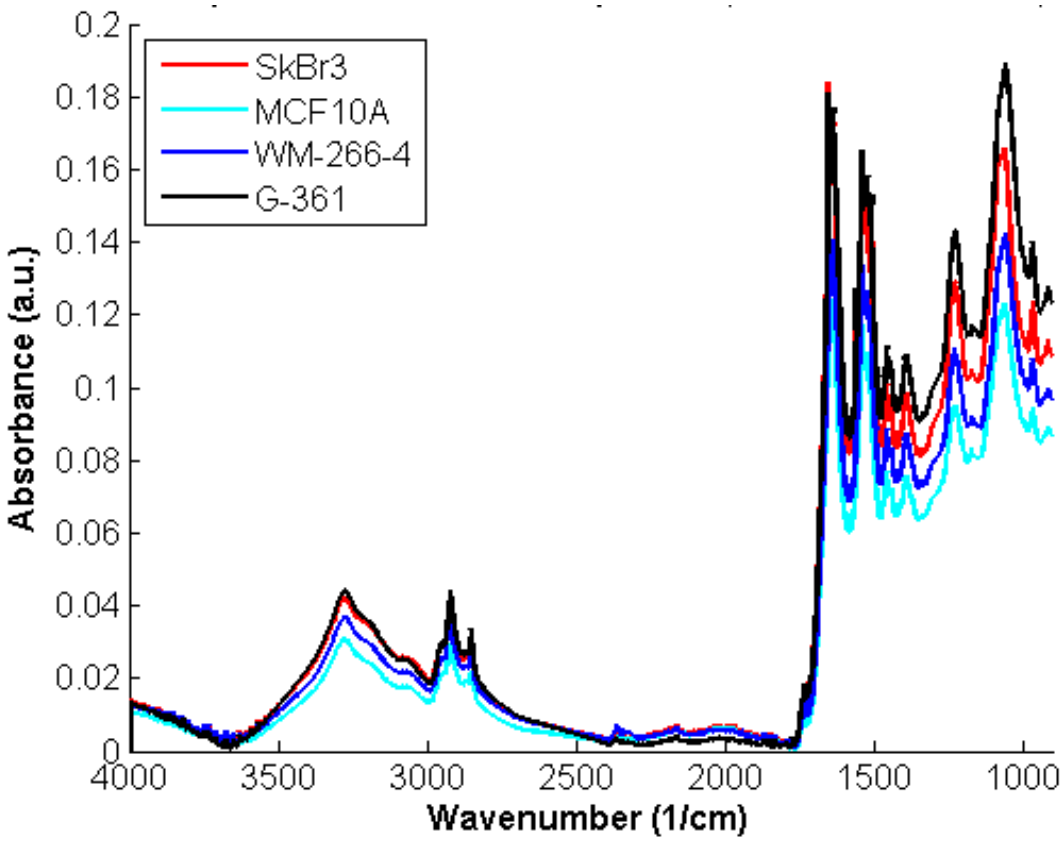



**Figure 1.1. This is a sample spectra depicting four cell lines (breast and skin, normal and cancer). The lipids are between 3000 and 1700 $cm^{-1}$, the proteins are between 1700 and 1400 $cm^{-1}$, and the nucleic acids (RNA, DNA) are between 1300 and 950 $cm^{-1}$.**

## 1.2. FTIR Spectroscopy Literature Review[4-43]

A large amount of data is available in the area of utilizing FTIR spectroscopy for mid-IR cancer detection. Because of what was learned from our research, and due to the initial readings of academic papers on the subject, we decided a literature review should be performed on this topic to gather a better understanding of what data was already available and what standard techniques existed.

Initially 72 academic papers were gathered with a cursory review; from those, 40 were selected for a more detailed review. Of these 40 papers, the following are the percentages of papers that dealt with specific cancers: 22% breast, 20% cervical, 13% lung, 11% skin, 7% prostate, 4% gastric, 2% endometrial, 2% esophageal, 2% leukemia, and 2% lymphoid.

We were specifically looking to categorize two areas of this research, sample preparation techniques and data acquisition techniques. The following sections will depict the data we retrieved from the literature review on these 40 papers.

### 1.2.1. Sample Preparation Methods

During the literature review, we were looking for the following three areas of sample preparation techniques: whether or not the samples were fixed, frozen, or neither; how the samples were dried; and how the samples were rinsed.

#### 1.2.1.1. Fixed, Frozen, or Neither

The researchers in 28% of the papers stated they fixed their samples, and 11% of the papers stated they froze their samples. However, 61% of the papers did not define either way. A large percentage of the authors that did not define their process most likely neither fixed nor froze their samples, although the authors never explicitly stated.

#### 1.2.1.2. Dried

The most common drying method was to air dry the samples; 35% of the authors stated they used this method to prepare their samples. 13% of the authors used nitrogen gas to dry their samples, and 7% used other methods of drying, which ranged from carbon dioxide gas to baking the samples in an oven at 50° C. Surprisingly, 46% of the researchers did not define how they dried their samples. This group of papers either never stated that they dried their samples or stated they dried the samples but never specified the drying process used.

#### 1.2.1.3. Rinsed

The most common form of rinsing the samples was to use a phosphate buffered saline (PBS) solution; 23% of the authors used this rinsing solution. 15% of the authors used DI water which, as previously mentioned, we proved to be the superior method. 23% of the papers stated they used other methods, such as a 0.9% sodium chloride solution, 0.4% potassium chloride solution, hexane, or undefined water. (The unspecified water was placed in this group because it was not mentioned if it was DI, sterilized, or tap water.) Finally, 40% of the papers did not state what was used to rinse the samples. This group either did not state that they rinsed the samples or stated they did rinse the samples but never specified what was used to perform the rinsing.

### 1.2.2. Data Acquisition Methods

#### 1.2.2.1. Analysis Techniques

In the literature review, the following two main analysis techniques were utilized by researchers: analyzing and comparing peak intensity and analyzing and comparing peak frequency shifts.

##### 1.2.2.1.1. Peak Intensity

The most common technique by far was to analyze the peak intensities. The authors that utilized this technique used several methods to analyze the peak intensities, such as comparing the intensity of specific wavenumbers between cancer and non-cancer samples. This method was the most simple and tended to provide the least-informative solution. The most common analysis method was to compare the ratios of various peaks. The most common ratio used was the two protein peaks Amide I (~1650 $cm^{-1}$) and Amide II (~1535 $cm^{-1}$); some authors used several peaks in their ratio calculations. Several of the researchers also took the mean average of the spectrum or a spectral integral over a range. The range most frequently used tended to be between 1800 $cm^{-1}$ and 900 $cm^{-1}$.

##### 1.2.2.1.2. Frequency Shift

The second technique used by researchers in this literary research was to compare frequency shifts of specific peaks; this method was not a common technique. However, there seemed to be some questions with this analysis method

that require more investigation. Mainly, the frequency shifts seen between cancer and normal samples seem to be smaller than the frequency range recorded in the 40 papers that define specific peaks.

#### 1.2.2.2. Attenuated Total Reflectance

The majority of researchers that use FTIR spectroscopy for cancer detection also incorporate an ATR attachment in their data collection. One issue with the ATR, however, is that the ATR pressure on the sample does affect the spectral intensity of the sample. In other words, adjusting the pressure on the sample affects the intensity of the signal. This issue was not addressed by any of the papers other than to state that "gentle pressure" was applied with the ATR. In order to address this issue, the majority of researchers normalize their data to a specific wavenumber. See Figure 1.2 for a definition of the various categories and wavenumber definitions.[20] The four more common frequencies used for normalization were Amide I (1653), Amide II (1545), phosphate stretching (1237), and glycine stretching (2927). Special care should be noted to observe the two columns depicting the frequency for various peaks of normal and cancerous samples. For example, the frequency shift for the Amide I peak between normal and cancer samples is only 3 $cm^{-1}$; the shift for Amide II is only 1 $cm^{-1}$. However, in the 40 paper review, the authors defined the Amide I peak anywhere between 1641 and 1658 $cm^{-1}$. Likewise, Amide II was categorized between 1533 and 1545 $cm^{-1}$. This area of concern requires further investigation before the frequency shift analysis method can be validated by my research.

| Normal ($cm^{-1}$) | Malignant ($cm^{-1}$) | Vibrational mode and main contribution |
|---|---|---|
| 933 | 923 | – |
| 964 | 967 | $\nu_s PO_4^{=}$ of phosphorylated proteins and nucleic acid |
| 995 | 999 | $\nu_s PO_4^{=}$ of RNA |
| 1024 | 1021 | $\nu$C–O + $\delta$C–O of glycogen |
| 1049 | 1057 | $\nu$C–O + $\delta$C–O of glycogen |
| 1080 | 1086 | $\nu_s PO_2^-$ of nucleic acids |
| 1155 | 1171 | $\nu$C–O of side chain groups in proteins |
| 1237 | 1239 | $\nu_{as} PO_2^-$ of nucleic acids |
| 1301 | 1303 | – |
| 1340 | 1341 | – |
| 1399 | 1398 | $\delta_s CH_3$ of proteins and lipids and $\nu_s COO^-$ of fatty acids |
| 1453 | 1455 | $\delta_{as}$ $CH_3$ of proteins and lipids |
| 1465 | 1466 | $\delta CH_2$ of lipids |
| 1545 | 1544 | Amide II of proteins |
| 1653 | 1656 | Amide I of proteins |
| 1741 | 1740 | $\nu$C=O of lipids |
| 2853 | 2853 | $\nu_s CH_2$ of lipids |
| 2927 | 2924 | $\nu_{as} CH_2$ of lipids |
| 2957 | 2957 | $\nu_{as} CH_3$ of lipids, nucleic acids and proteins |
| 3291 | 3294 | Amide A of proteins and $\nu$O–H of carbohydrates |

**Figure 1.2. Frequency ($cm^{-1}$) and assignments of the major bands in the infrared spectra of normal and malignant tissues[20]**

### 1.2.2.3. Normalized Data

The most common peak the authors used to normalize their data was Amide I; 24% of the researchers used this peak. While only 3% of the papers used Amide II as the normalizing factor, 5-6% of the papers used the phosphate stretching frequency (~1240 $cm^{-1}$). Likewise, 5-6% of the papers used the glycine stretching

frequency (~2920 $cm^{-1}$). 19% of the papers actually normalized the data to ratios, the most common ratio being Amide I / Amide II. The papers that normalized the data to ratios typically did not contain much detail other than to possibly mention which peaks were used in the ratio.

10% of the researchers normalized their data to a spectral integral over a specified frequency range. The range typically covered 1800 $cm^{-1}$ to 900 $cm^{-1}$. 7% of the authors used other peaks to normalize their data, such as C-O stretching (1024 $cm^{-1}$). Finally, 22% of the papers did not define which peak was used to normalize their data. This category also included the authors that did not state whether or not they actually did normalize the data. One of the papers actually specifically stated they did not normalize their data, although no reason was given for not taking this approach and no technique was mentioned for counteracting the ATR pressure variation affecting the signal intensity.

### 1.2.3. Literary Review Discussion

Generally, the majority of the authors scrutinized in this review did not justify why a specific technique was used; no motive was given. In fact, there is a general vagueness in academic papers dealing with FTIR spectroscopy utilized for cancer detection. This fact is depicted in the large percentage of papers that were "not defined" for techniques such as sample preparation and data acquisition techniques. Papers would often state that a specific task was performed, but rarely did the paper state why that technique was used. This ambiguity in the field can

possibly be attributed to a lack of standards in the field, especially for sample preparation techniques and data collection methods.

Included in this vagueness is a general lack of justification for why specific peaks were used to normalize the data. However, a very small percentage of papers actually did justify why they normalized their data to a wavenumber. The most common peak utilized, Amide I, was typically justified simply because it is normally the highest intensity in the spectrum. However, several of the papers that used the phosphate stretching frequency specifically stated why they did *not* use Amide I, which was because it is much more unstable that other frequencies. The Amide I signal is much more sensitive to environmental factors and would often have significant peak shifts. The reason these papers used the phosphate stretching peak is because it is very stable and virtually unaffected by environmental factors such as washing or the creation of a dry film on the slides.

### 1.2.4. Literary Review Conclusion

Generally, no paper in the literary review dealt with issues of varying ATR pressure changing the spectra's signal peak intensity. Although most researchers normalize their data, no standards seem to exist as to which frequency peak or spectral region should be used. Likewise, no standards seem to exist for sample preparation techniques, although clearly the sample preparation method plays an important role in obtaining a clean spectral signal of the sample. A large percentage of papers do not adequately define the techniques used, which could be partially due to a lack of standards in the FTIR cancer detection processes.

Standards are needed in sample preparation and data collection to help validate FTIR spectroscopy as a viable cancer detection technique.

Chapter 2

# Mid-IR Spectral Investigation of Normal and Malignant Cervical Tissue Samples

Cervical cancer used to be the leading cause of cancer deaths for women in the United States. A significant decrease has been associated with most women in the United States obtaining regular pap tests to identify cervical precancer before it can actually develop into cancer. However, quick, easy, and accurate detection in cervical cancer and precancer can still reduce medical costs, time between biopsy and patient results and treatment, and assist with automated prescreening of tissue readings. In addition, according to the World Health Organization, cancer is the leading cause of death worldwide, and cervical cancer still accounts for approximately 20% of all cancer mortality rates in women in low and middle-class populations.[44-48]

In this study, we applied Daylight Solution's quantum cascade laser-based microscope (Spero) to investigate mid-infrared characteristics of normal and

malignant cervical tissue. First, sample thickness was evaluated by sending various thickness samples to Daylight Solutions for evaluation. Second, samples fixed in paraffin were evaluated as is and de-paraffinized. Finally, the normal and malignant samples were prepared and evaluated utilizing the unique capabilities of Spero.

## 2.1. Materials

### 2.1.1. Spero[49-51]

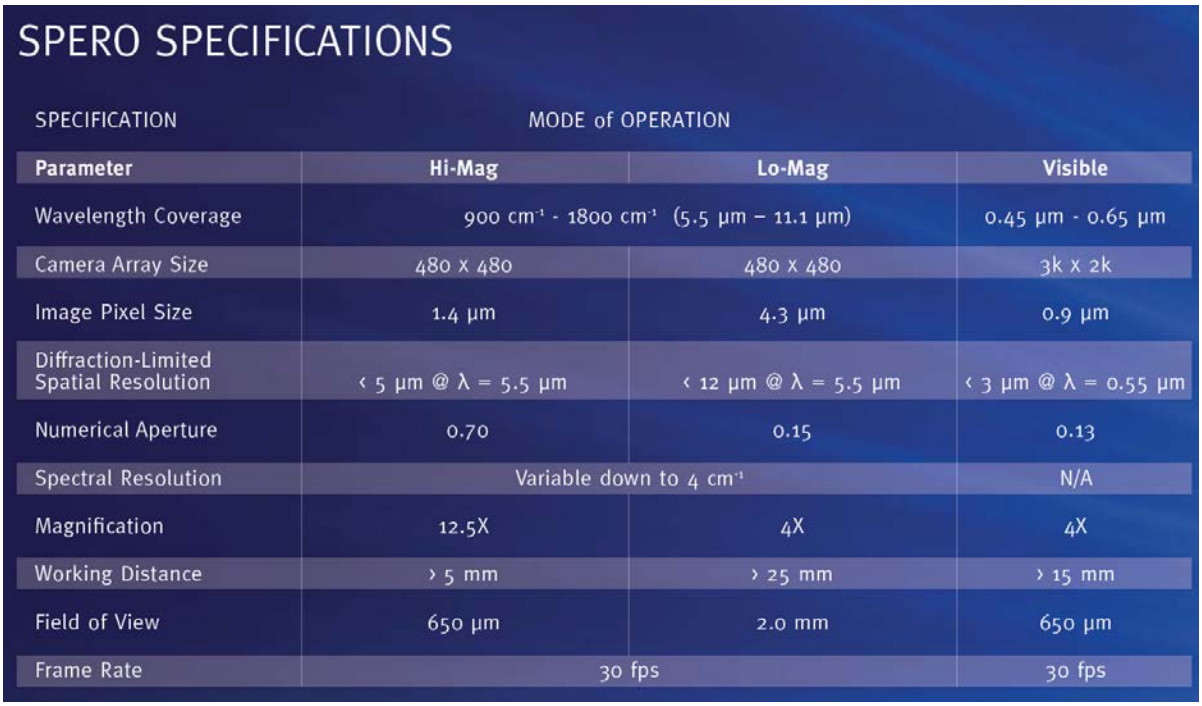
Daylight Solutions is a company based out of San Diego, California, and they have developed the first laser-based infrared microscope, Spero (see Figure 2.1)[49]. The capabilities of this microscope include both high-resolution imaging over a wide field of view and rapid live mode imaging at 30 frames per second. It has potential in a wide range of fields, including cancer diagnostics, drug discovery, materials research, forensics, and semi-conductor analysis. However, we have focused on the cancer diagnostic properties of this instrument.


**Figure 2.1. Daylight Solutions' quantum cascade laser-based microscope, Spero.[49]**

Spero is an infrared hyperspectral imaging microscopy platform based around tunable quantum cascade laser technology. The quantum cascade lasers are coupled with a series of mid-IR refractive objectives and an uncooled microbolometer camera. The base of the microscope seen in Figure 2.1 holds the quantum cascade lasers, which has large spectral coverage from 1800-900 $cm^{-1}$, or 5.5-11.1 μm. Spero contains four lasers that have some overlap, allowing the large spectral coverage. The uniqueness of Spero stems from the QCLs, which is tunable,

allowing the samples to be illuminated with discrete frequencies across the mid-IR fingerprint region.

Spero can be used in transmission or reflection modes and also includes live imaging at discrete mid-IR wavelengths. In addition, data can be collected at specified discrete wavelengths instead of covering the entire spectrum, which turns the data acquisition process into seconds. The live mode can utilize either a high or low magnification camera. The high magnification camera enables a live-field view of 650 μm x 650 μm at 1.34 micron pixel resolution with a 0.7 numerical aperture and a 4.3 μm spatial resolution. The low magnification camera has a 2000 μm x 2000 μm at 4.25 micron pixel resolution field-of-view with a 0.15 numerical aperture and a 25 μm spatial resolution. See Figure 2.2 for specifications of the instrument's three modes of operation.[51] Because Spero uses a microbolometer in place of the FTIR's MCT detector, liquid nitrogen is not needed. The instrument's optics are refractive, making the numerical aperture much higher than instruments utilizing the typical reflective optics.



SPERO SPECIFICATIONS

| SPECIFICATION | MODE of OPERATION | | |
|---|---|---|---|
| **Parameter** | **Hi-Mag** | **Lo-Mag** | **Visible** |
| Wavelength Coverage | 900 $cm^{-1}$ - 1800 $cm^{-1}$ (5.5 µm – 11.1 µm) | | 0.45 µm - 0.65 µm |
| Camera Array Size | 480 x 480 | 480 x 480 | 3k x 2k |
| Image Pixel Size | 1.4 µm | 4.3 µm | 0.9 µm |
| Diffraction-Limited Spatial Resolution | ‹ 5 µm @ λ = 5.5 µm | ‹ 12 µm @ λ = 5.5 µm | ‹ 3 µm @ λ = 0.55 µm |
| Numerical Aperture | 0.70 | 0.15 | 0.13 |
| Spectral Resolution | Variable down to 4 $cm^{-1}$ | | N/A |
| Magnification | 12.5X | 4X | 4X |
| Working Distance | › 5 mm | › 25 mm | › 15 mm |
| Field of View | 650 µm | 2.0 mm | 650 µm |
| Frame Rate | 30 fps | | 30 fps |

**Figure 2.2. Spero specifications of high magnification, low magnification, and visible modes of operations.[51]**

A few comparisons specifically outlining key advantages of FTIR include far superior spatial resolution which allows an increased field of view and the speed at which data can be acquired, whether the full spectrum is collected or discrete wavelengths. In addition, one of the greatest advantages is the live discrete video and imaging capabilities, which allows the researcher to accurately select the point of data collection on the sample in real-time.

### 2.1.2. Tissue Samples[52]

We acquired all our cervical tissue samples, both normal and malignant, from the Cooperative Human Tissue Network (CHTN) based out of the University of

Alabama Birmingham. Table 2.1 and Table 2.2 show additional histological information provided by CHTN with each cervical sample. The normal tissue is categorized by two methods, ethnicity and the cervical location the sample was removed. Because CHTN classified the ethnicity of all the samples as Black, White, Asian, or Hispanic, we used these same classifications throughout our documentation. The malignant tissue samples did not include the location of the sample.

| | | |
|---|---|---|
| Cervical Location | Ectocervix | 7 |
| | Endocervix | 8 |
| | Ecto/Endocervical Junction | 1 |
| | Unknown | 4 |
| Ethnicity | White | 14 |
| | Black | 5 |
| | Hispanic | 1 |

Total Normal Cervical Samples: 20

**Table 2.1. Normal cervical samples.**

| | | |
|---|---|---|
| Ethnicity | White | 8 |
| | Black | 4 |
| | Asian | 1 |
| | Unknown | 1 |

Total Cervical Cancer Samples: 14

**Table 2.2. Malignant cervical samples.**

The tissue samples included sparse diagnostics and comments received from the medical staff and principle investigators cooperating with CHTN. The specimens went through an initial diagnostic examination by a trained pathologist, which were then verified through a frozen section and evaluation of permanent histopathology. A total of twenty normal cervical specimens and fourteen malignant cervical specimens were received from CHTN.

### 2.1.3. Optical Slides and Windows

Standard glass slides were used for all hematoxylin and eosin (H & E stain) samples. Calcium fluoride (CaF2) windows from International Crystal Laboratories (ICL) were selected for transmission data collection due to the characteristics that allow transmission through the MIR region, as depicted in Figure 2.3.[53] CaF2 windows were also used by Daylight Solutions in their data collection and recommended by them in order for consistency.

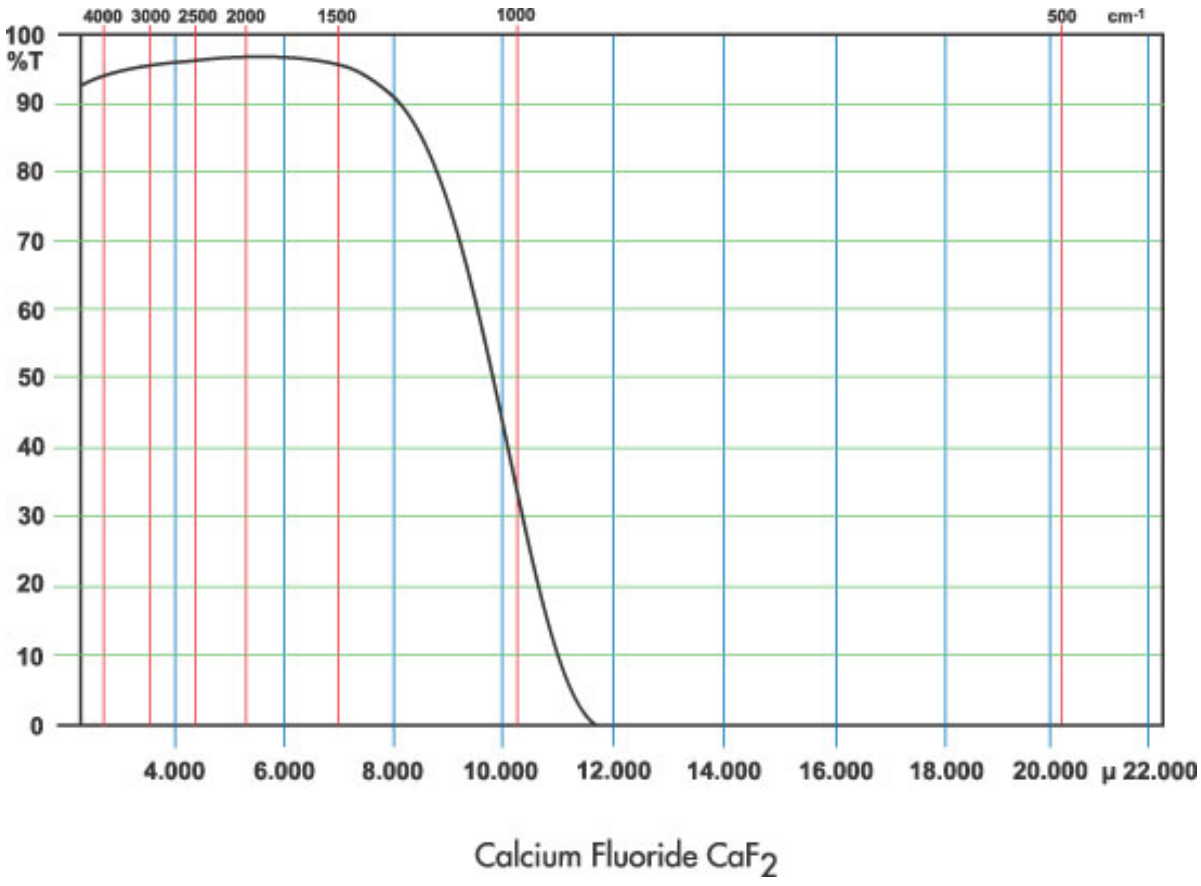



**Figure 2.3. Optical properties of calcium fluoride optical crystals.[53]**

### 2.1.4. Biology of Cervical Tissue Specimens

The cervix is the lower part of the uterus; it is the connection between the uterus and the birth canal (see Figure 2.4)[54]. The ectocervix is the lower portion of the cervix that is exposed to the birth canal. It is made of both squamous and columnar epithelial cells. The endocervix is the upper portion of the cervix that opens to the lower uterus; it is made up of columnar epithelial cells. The majority of cancer occurs in either the ectocervix or the endocervix; 80-90% of cervical cancer is squamous cell carcinoma.[54]

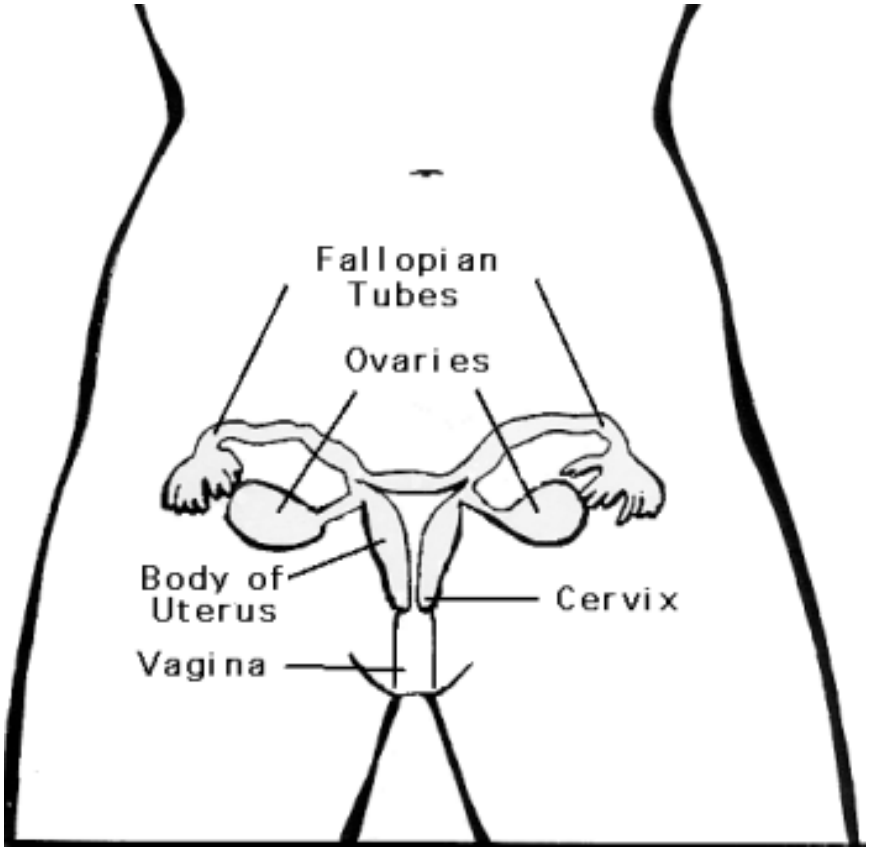



**Figure 2.4. Location of cervix.[54]**

The key area of interest on the cervical specimens is the epithelium and its subsequent layers as depicted in Figure 2.5.[58]

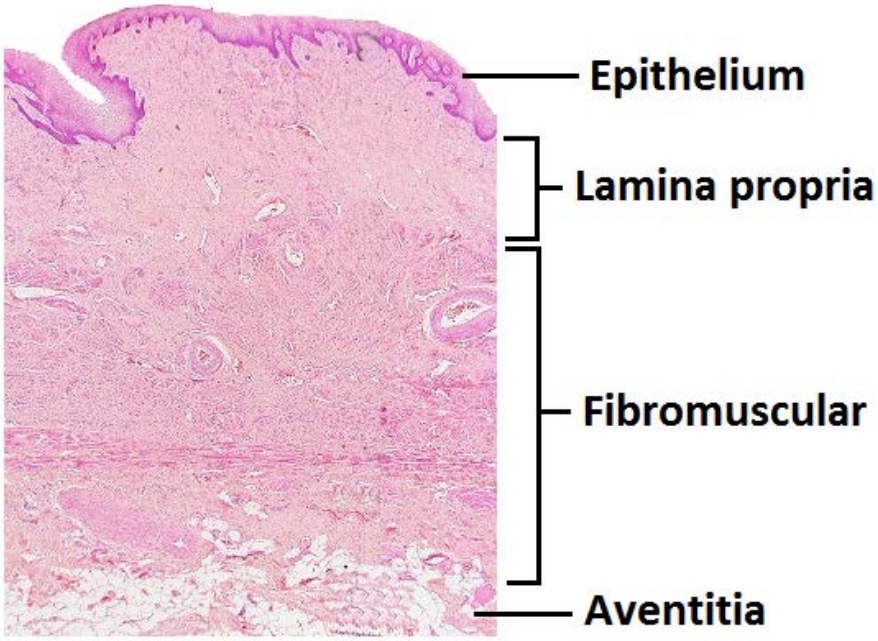



**Figure 2.5. Cervical tissue layers.[58]**

Within the epithelium, both columnar and stratified squamous cells can be present (see Figure 2.6)[59].

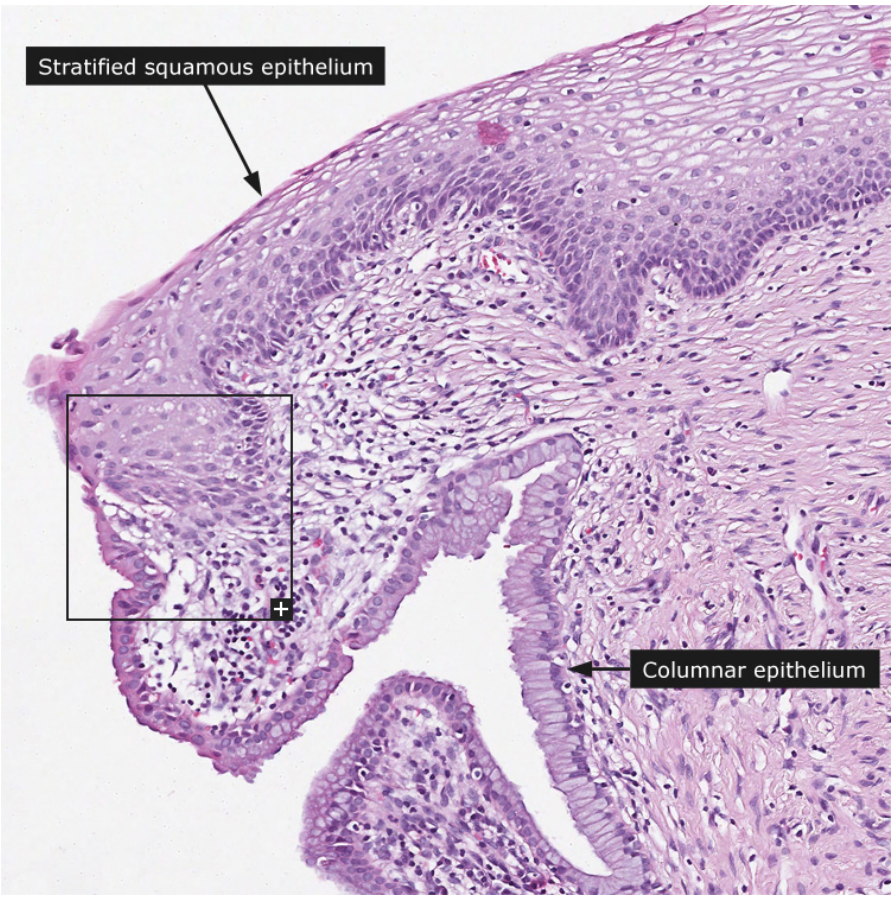



**Figure 2.6. Stratified squamous and columnar epithelium.**[59]

The thin layer of cells separating the epithelium from the lamina propria is the basal layer, which was also a focus of data collection (see Figure 2.7)[60]. The flat cells depicted in Figure 2.7 are also a good example of stratified squamous epithelium cells.

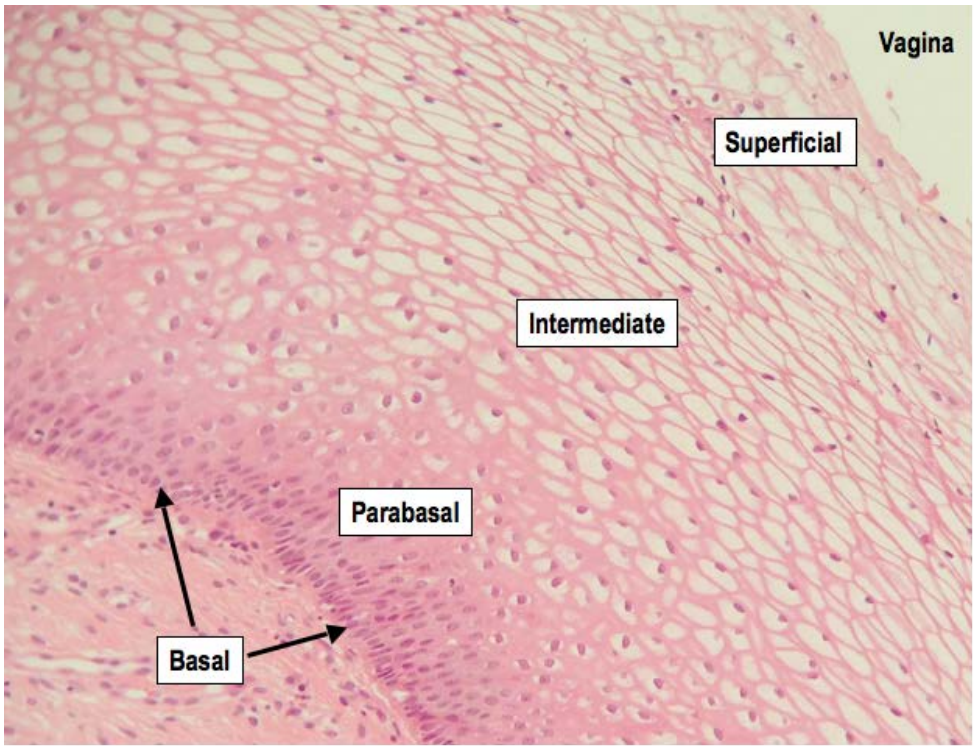



**Figure 2.7. Epithelium layers, basal layer.[60]**

## 2.2. Methods

### 2.2.1. Sample Preparation

The cervical tissue samples arrived frozen from CHTN in dry ice. Baylor College of Medicine (BCM) Pathology and Core Lab was utilized to prepare the tissue samples. The tissue specimens were all fixed and embedded in paraffin.

Initially, the following sample thicknesses were excised from a cervical specimen: 5, 7, 10, 12, and 15 µm. These samples were then sent to Daylight Solutions in San Diego, California, to evaluate optimal sample thickness for transmission data acquisition on Spero. Daylight Solutions evaluated the various

thicknesses and concluded the optimal sample to be 5-7 μm thick. They also concluded the optimal thickness of the CaF2 windows to be 2 mm and provided a holder for ICL's 25 mm diameter disc.

BCM's Pathology and Core Lab fixed each specimen in formalin; the specimens were then each embedded into a block of paraffin. Two tissue samples, approximately 5-7 μm thick, were excised from each specimen; one sample was placed on a standard glass slide and one sample was placed on a 2 mm thick calcium fluoride (CaF2) window from ICL. The samples were then de-paraffinized to prevent artifacts in the spectra.

Samples on the standard glass slides were stained with hematoxylin and eosin (H & E stain), as seen in Figure 2.8. These were used to identify regions of interest on the CaF2 samples (see Figure 2.9), which were used for transmission data collection on Spero. All samples were then stored in a desiccator while not in use.

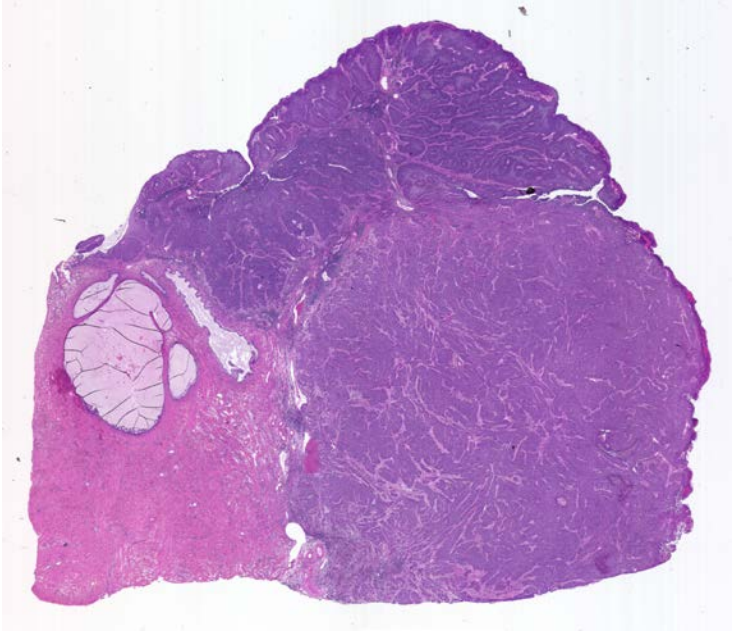


**Figure 2.8. Example of H&E stain, malignant cervical specimen 03-04-A370a.**

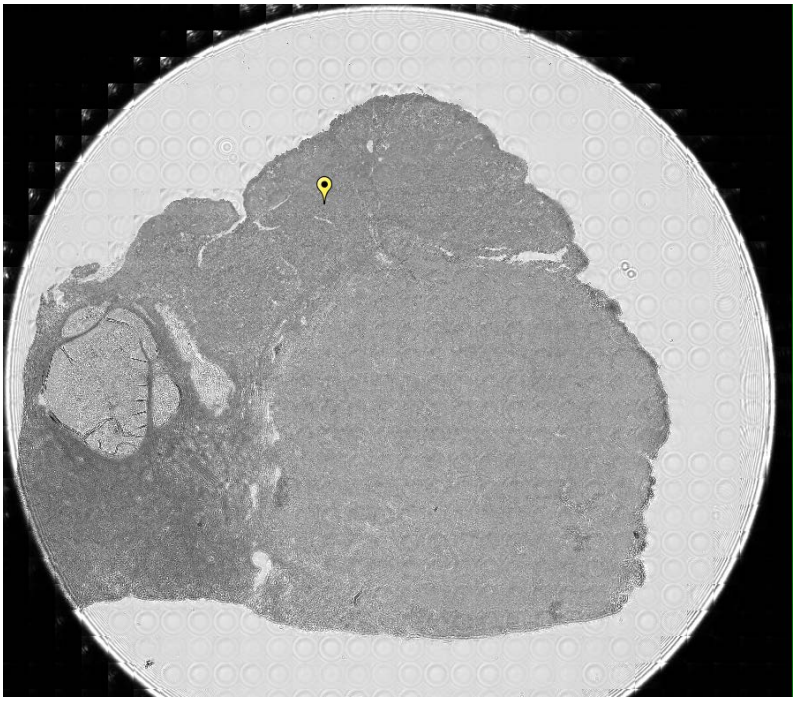


**Figure 2.9. Example of Spero image, malignant cervical specimen 03-04-A370a.**

### 2.2.2. Data Acquisition Process

The Spero microscope was set up on a floating optical table to absorb environmental vibrations and other noise interference. A constant flow of nitrogen gas at approximately 5 cc/second was required to purge the sample stage enclosure of water vapor and other contaminants. Initially we obtained nitrogen tanks to perform the purge, but the tanks only lasted 1-2 days of continuous 5 cc/second flow. We then changed to the facility's nitrogen ports available in the lab by running approximately 80 feet of hosing to the nearest nitrogen port.

Per recommendations by Daylight Solutions, we ensured the microscope's lasers had sufficiently warmed up by having the instrument on for at least an hour prior to any use. The back of each 25 mm diameter CaF2 window was gently cleaned with optical tissue and an ethanol solution to remove dust particles. The sample was placed on a holder provided by Daylight Solutions and positioned on the stage. The stage cover was then closed and the system was allowed to purge with nitrogen for at least five minutes prior to any data or background collection. A new background was taken on a clean portion of the same slide the sample was on prior to any data collection of a new sample; a new background was taken at a minimum every three data points of that sample or every 45 minutes.

Each Spero image was attained using the high-magnification infrared objective. The tissue sample was oriented similarly to the H&E stained image, as seen in Figure 2.8 and Figure 2.9. Each sample on a CaF2 window was compared to the H&E stain to identify regions of interest to collect data points. The focus of data

collection on the normal specimens was on the epithelium, basal, and lamina propria (basement membrane) and the adjoining layers. The focus of the malignant specimens was on both the tumor and the transition tissue adjacent to the tumor. Because the epithelium and basal layers were not clearly defined on the tumors, all spectral data was collected from the lamina propria. Normal versus malignant comparison analysis was performed with the lamina propria spectra from both the normal and malignant samples.

### 2.2.3. Peaks Corresponding to Biology[79,80]

Specific wavenumbers and ranges of wavenumbers in the mid-infrared spectrum correspond to unique biological characteristics. Some of the key characteristics are depicted in Figure 2.10[55] and Figure 2.11[20].

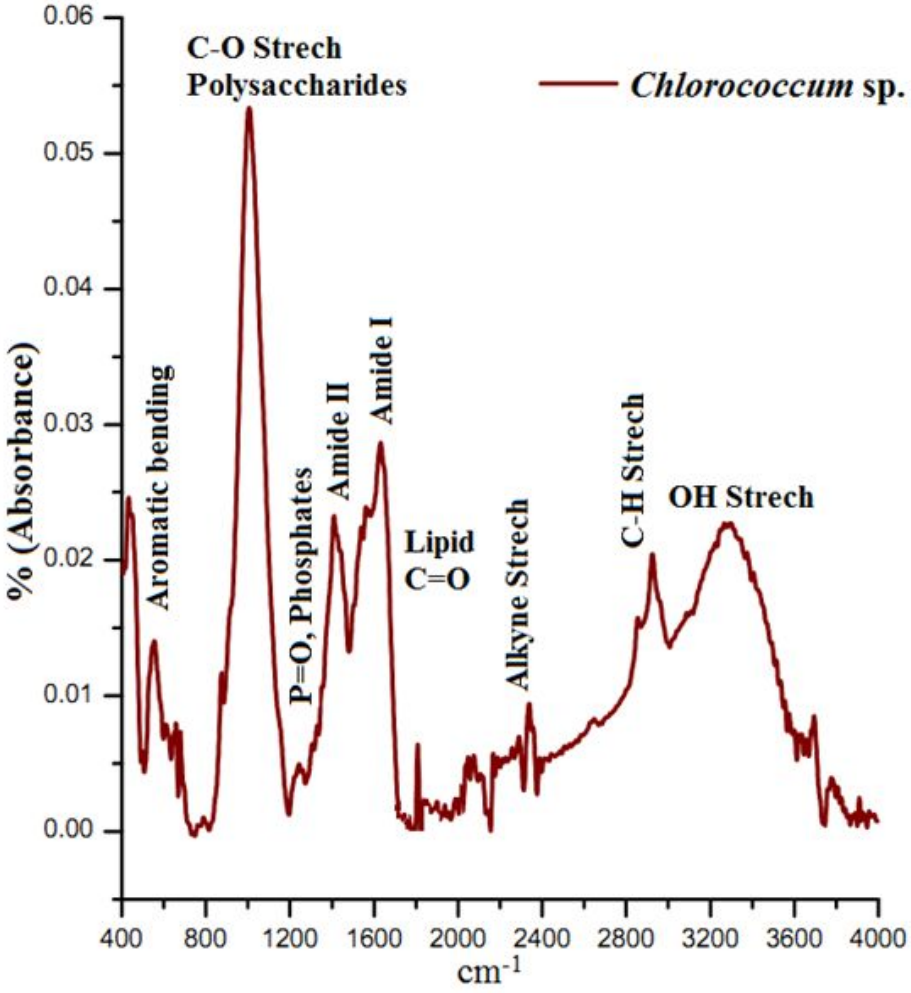



**Figure 2.10. Mid-IR spectrum peaks and corresponding cellular characteristics.**[55]

| Normal ($cm^{-1}$) | Malignant ($cm^{-1}$) | Vibrational mode and main contribution |
|---|---|---|
| 933 | 923 | – |
| 964 | 967 | $\nu_s PO_4^{=}$ of phosphorylated proteins and nucleic acid |
| 995 | 999 | $\nu_s PO_4^{=}$ of RNA |
| 1024 | 1021 | $\nu$C–O + $\delta$C–O of glycogen |
| 1049 | 1057 | $\nu$C–O + $\delta$C–O of glycogen |
| 1080 | 1086 | $\nu_s PO_2^-$ of nucleic acids |
| 1155 | 1171 | $\nu$C–O of side chain groups in proteins |
| 1237 | 1239 | $\nu_{as} PO_2^-$ of nucleic acids |
| 1301 | 1303 | – |
| 1340 | 1341 | – |
| 1399 | 1398 | $\delta_s CH_3$ of proteins and lipids and $\nu_s COO^-$ of fatty acids |
| 1453 | 1455 | $\delta_{as}$ $CH_3$ of proteins and lipids |
| 1465 | 1466 | $\delta CH_2$ of lipids |
| 1545 | 1544 | Amide II of proteins |
| 1653 | 1656 | Amide I of proteins |
| 1741 | 1740 | $\nu$C=O of lipids |
| 2853 | 2853 | $\nu_s CH_2$ of lipids |
| 2927 | 2924 | $\nu_{as} CH_2$ of lipids |
| 2957 | 2957 | $\nu_{as} CH_3$ of lipids, nucleic acids and proteins |
| 3291 | 3294 | Amide A of proteins and $\nu$O–H of carbohydrates |

**Figure 2.11. Frequency ($cm^{-1}$) and assignments of the major bands in the infrared spectra of normal and malignant tissues.[20]**

## 2.3. Results

Mid-IR spectral areas of the highest absorption in the cervical samples tend to be the two protein Amide peaks (1650 $cm^{-1}$ and 1545 $cm^{-1}$), along with phosphates (asymmetrical stretching at 1240 $cm^{-1}$ and symmetrical stretching at

1080 $cm^{-1}$). The epithelium layers also often express a glycogen absorption peak at 1025 $cm^{-1}$.

### 2.3.1. Normal Cervical Data Acquisition

Twenty normal cervical specimens were utilized for this experiment. Samples from each specimen were placed on glass slides and CaF2 windows; the samples on the glass slides were stained with H&E. Data for each sample on the CaF2 windows were collected on Spero and evaluated in this experiment; the locations of the data collection points were first identified on the samples stained with H&E. Each sample had multiple points of data collection; emphasis was placed on the epithelium and the adjacent layers. A few key normal cervical specimen examples are listed in this section. Refer to Appendix A for a complete listing of normal cervical specimens, including images from the H&E-stained slides, the images obtained through Spero's high-magnification objective, and the spectra from each data point collected.

#### 2.3.1.1. Specimen 04-10-A245a

Specimen 04-10-A245a is from a 32-year-old white female. The specimen is a 0.85 gram ectocervical sample. Figure 2.12 is an image of a sample of the specimen stained with H&E on a glass slide; a clear epithelium edge is apparent on the tissue sample, as depicted by the dark purple line along the edge of the tissue.

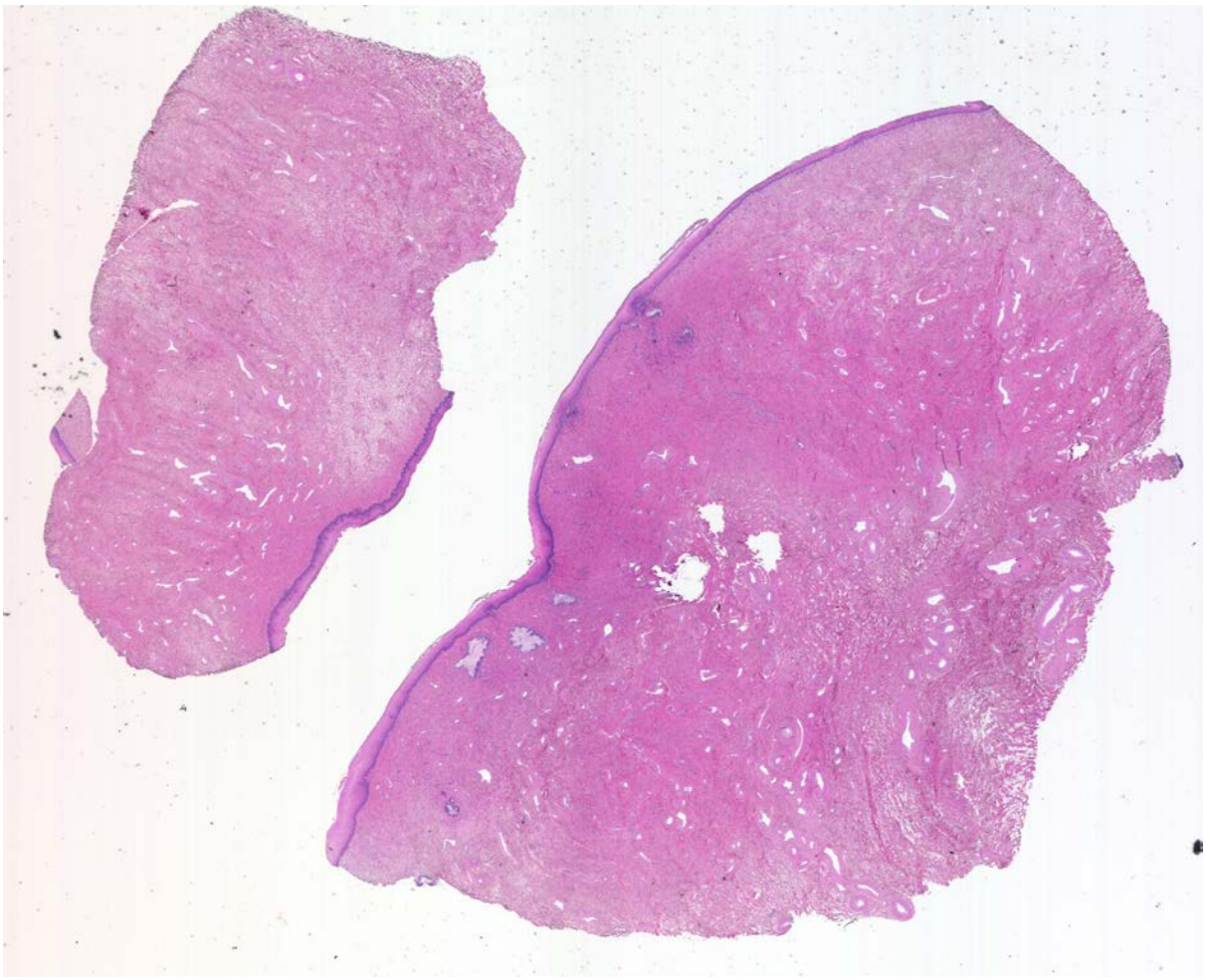


**Figure 2.12. Specimen 04-10-A245a, H&E stain, normal cervical tissue.**

Figure 2.13 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags J and K depict the data collection points on Sample 04-10-A245a, focusing on the epithelium and the nearby layers.

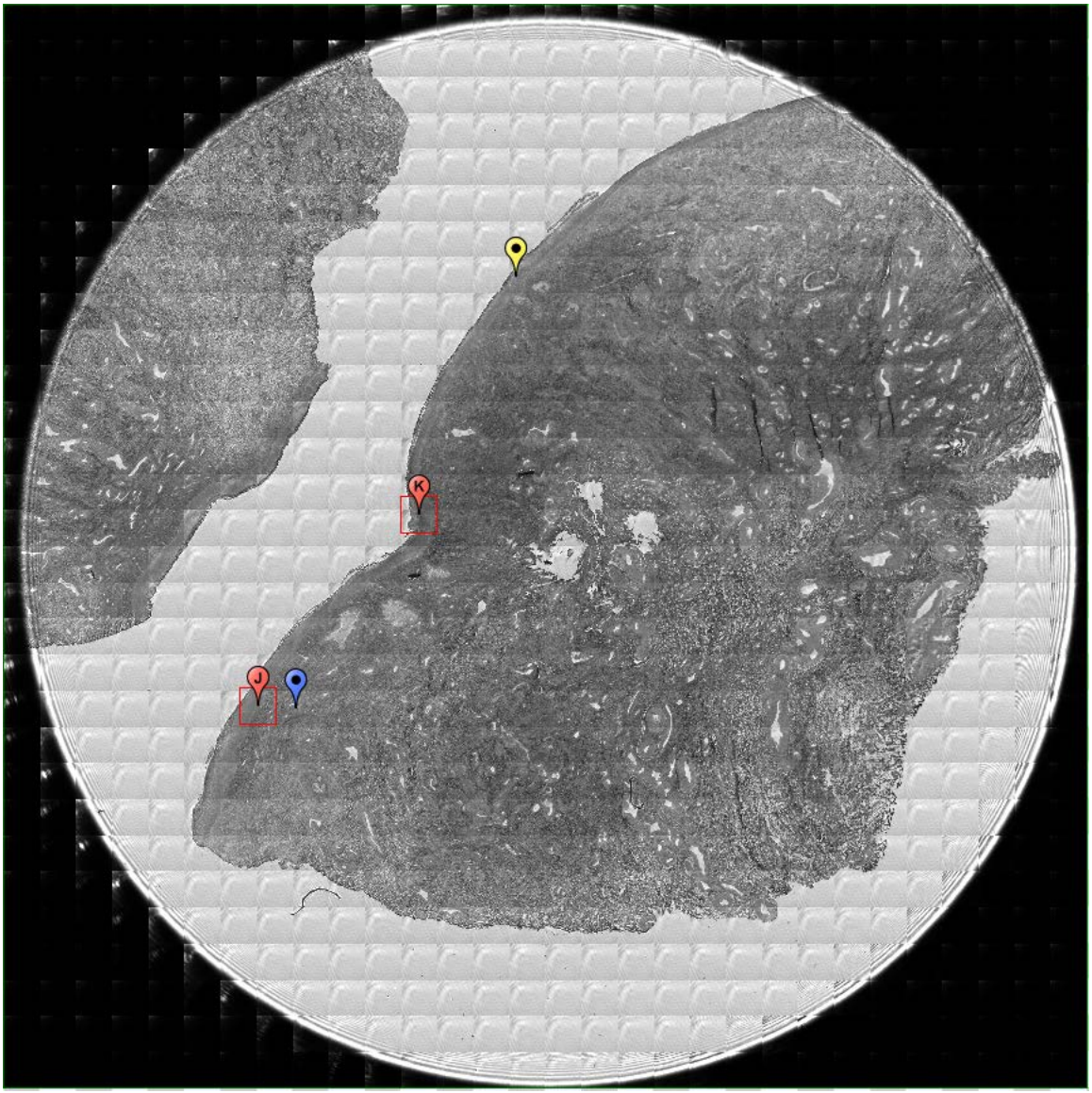



**Figure 2.13. Specimen 04-10-A245a, normal cervical tissue.**

#### 2.3.1.1.1. Sample 04-10-A245a, Point J

A 650 x 650 µm field-of-view image taken of the specimen at Point J (see Figure 2.14) was obtained from Spero's high-magnification objective. The epithelium is clearly evident in this figure, although the single-cell basal layer is not

as evident. The circular and flat squamous cells are also visible, especially near the birth canal.

Regions of interest were defined, which is where the spectra were actually taken on the sample. Region 1 is defining an area of interest on the stratified squamous epithelium; Region 2 is focusing on the basal and parabasal layer. Regions 3 and 4 are identifying regions of interest in the basement membrane and lamina propria.

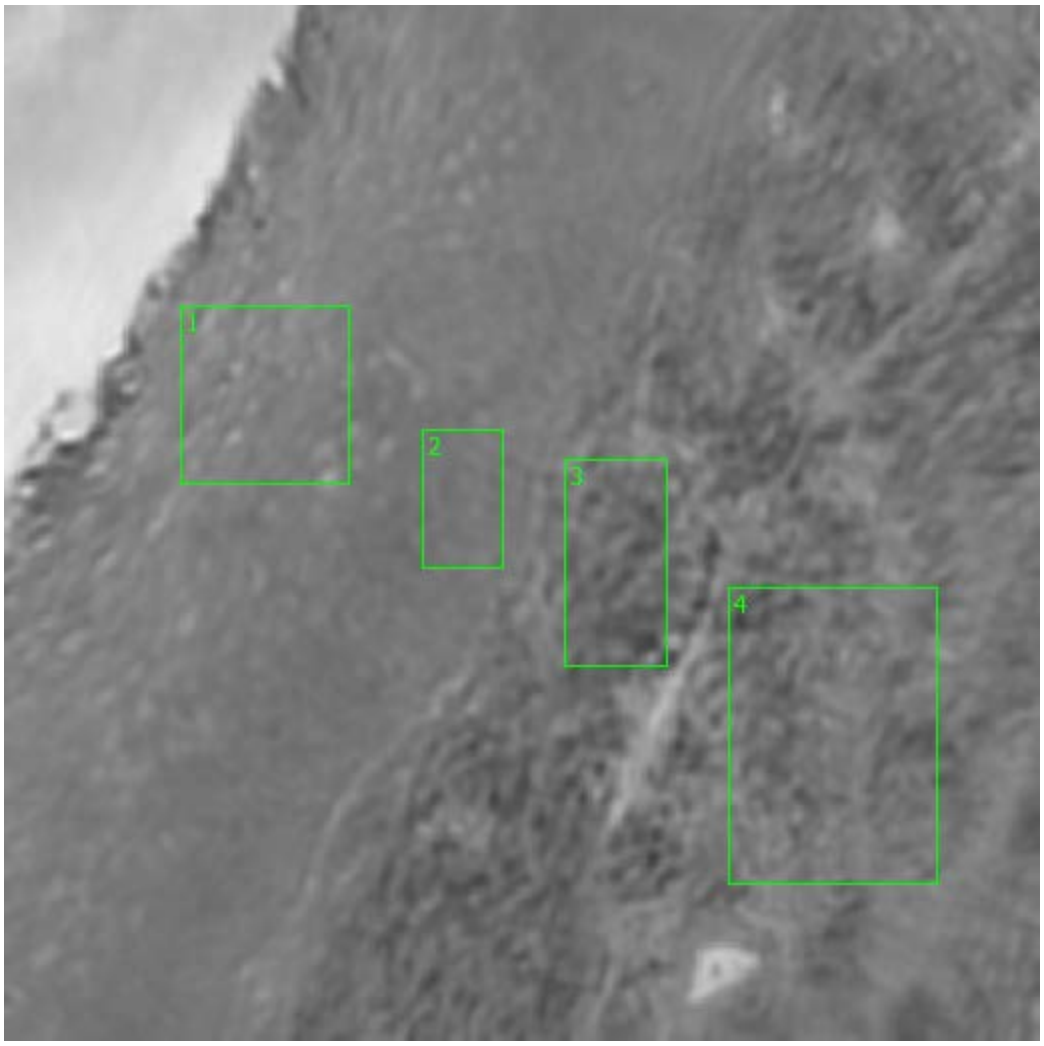



**Figure 2.14. Specimen 04-10-A245a, Point J Regions of Interest**
**Region 1: stratified squamous epithelium**
**Region 2: parabasal layer**
**Regions 3, 4: basement membrane, lamina propria**

The spectra from the four defined regions of interest on Point J are depicted in Figure 2.15.

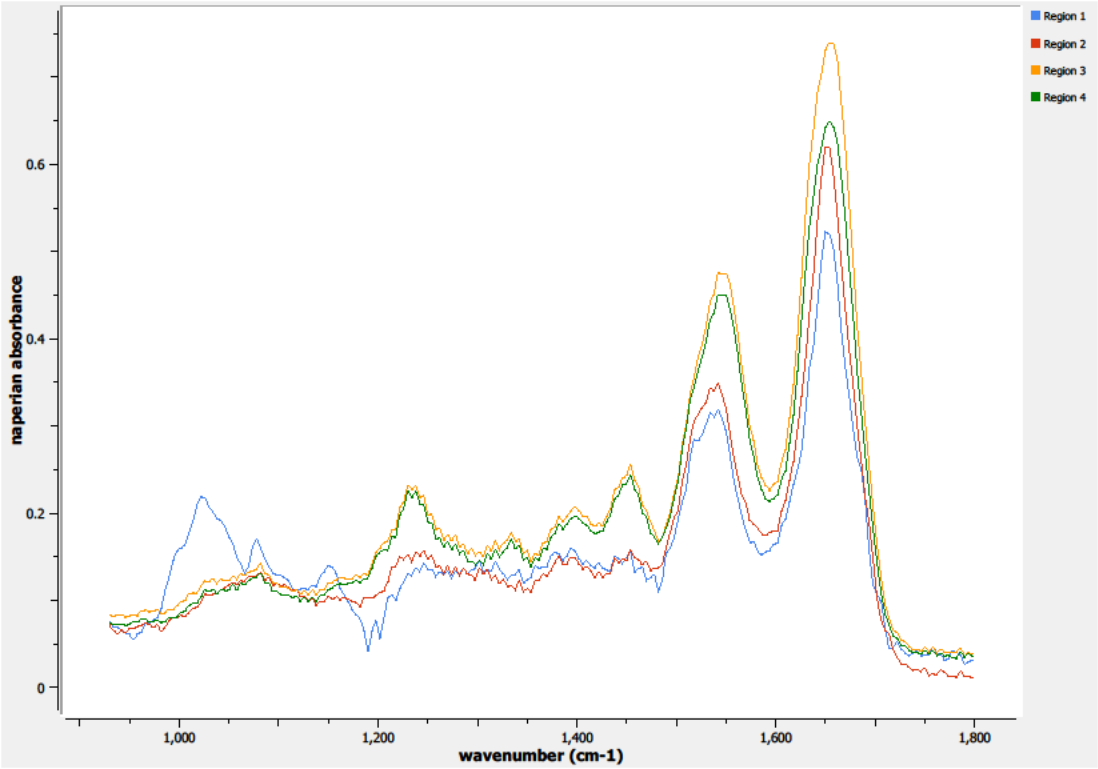



**Figure 2.15. Specimen 04-10-A245a, Point J Spectra**

#### 2.3.1.1.2. Sample 04-10-A245a, Point K

The high-magnification 650 x 650 µm field-of-view image was also taken at Point K (see Figure 2.16) on Sample 04-10-A245a. The epithelium again is clearly evident in this figure; the basal layer also appears to be present, which is why spectra were collected at Regions 4 and 5.

Flat squamous cells are evident in Region 2, and more rounded squamous and columnar cells are visible in the deeper tissue parabasal layer. Region 6 is defining a region of interest in the basement membrane or the lamina propria.

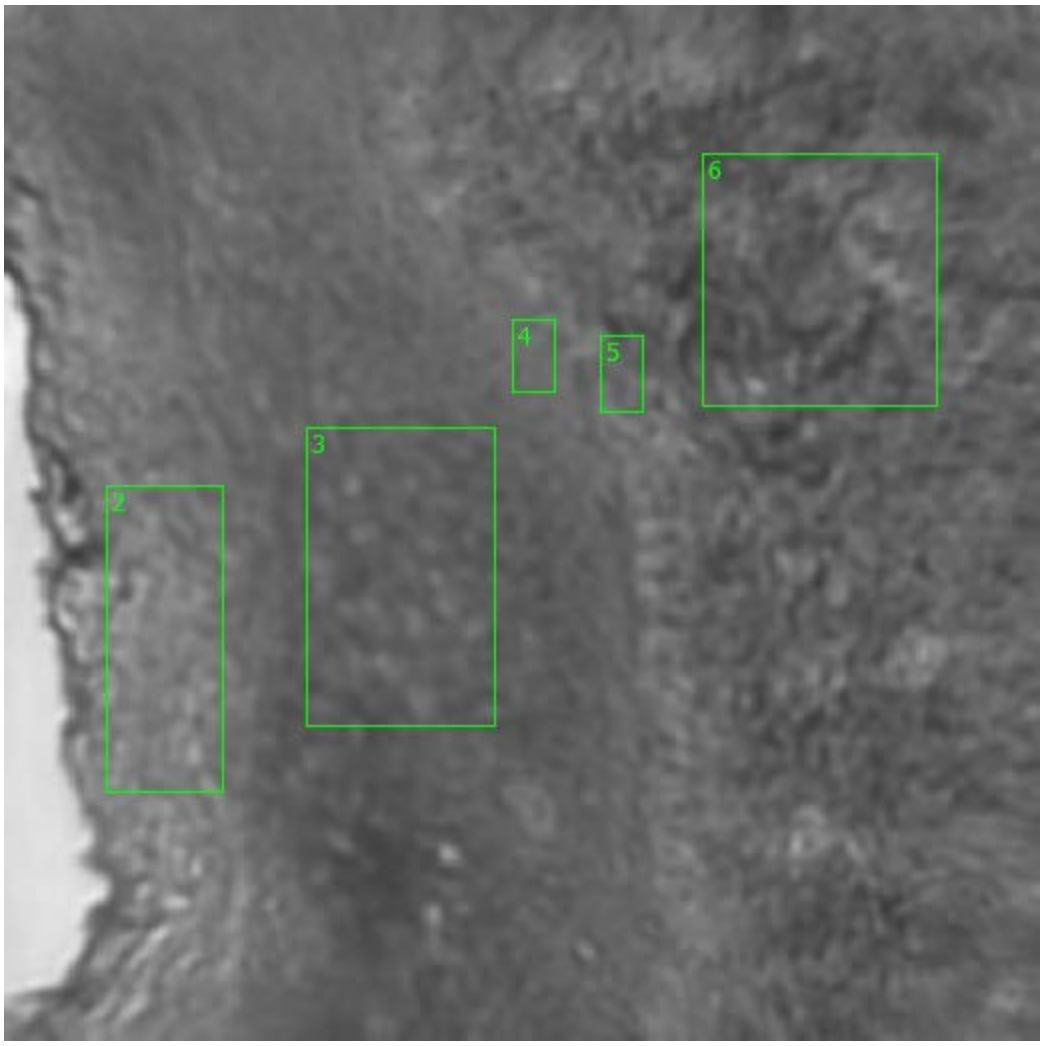



**Figure 2.16. Specimen 04-10-A245a, Point K Regions of Interest**
**Region 2: stratified squamous epithelium**
**Region 3: columnar epithelium**
**Region 4: basal layer**
**Regions 5, 6: basement membrane, lamina propria**

The spectra from the five defined regions of interest on Point K are depicted in Figure 2.17.

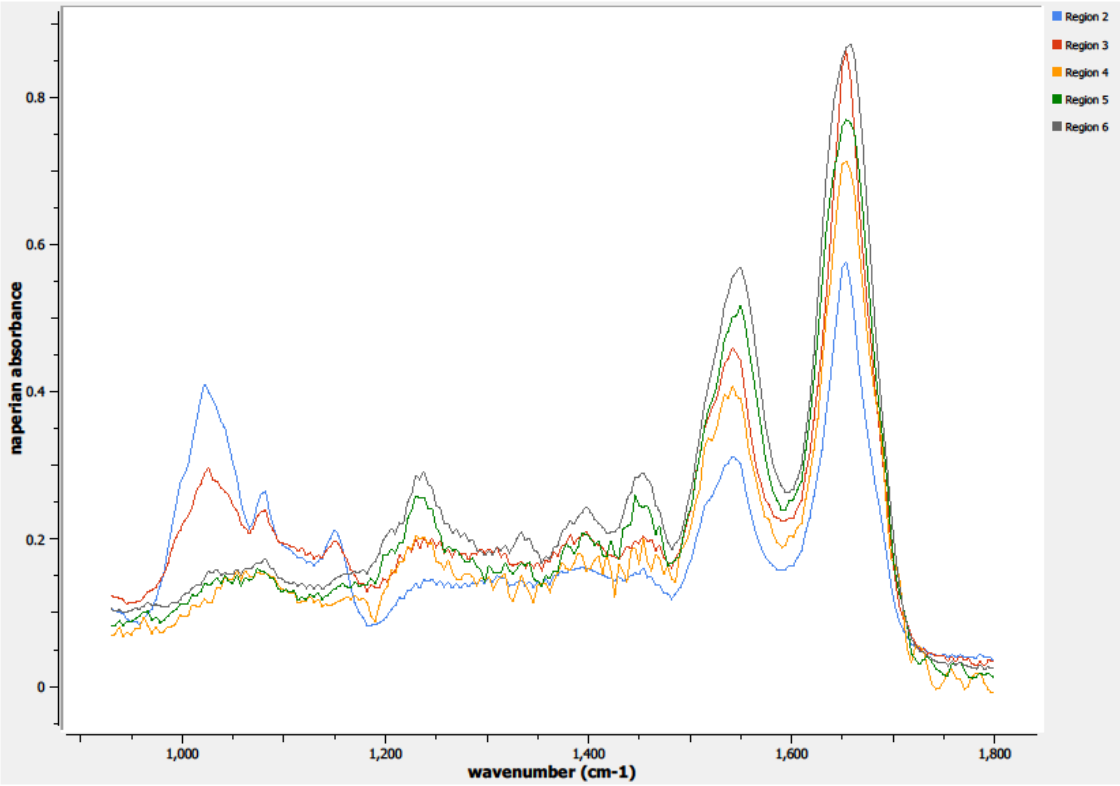



**Figure 2.17. Specimen 04-10-A245a, Point K Spectra**

### 2.3.1.2. Specimen 12-05-A046a

Specimen 12-05-A046a is from a 34-year-old white female. The specimen is a 1.36 gram endocervical sample. Figure 2.18 is an image of a sample of the specimen stained with H&E on a glass slide. The epithelium layer, although quite thin, is noticeably apparent on the tissue sample, as depicted by the dark purple line along the lower edge of the tissue in the figure.

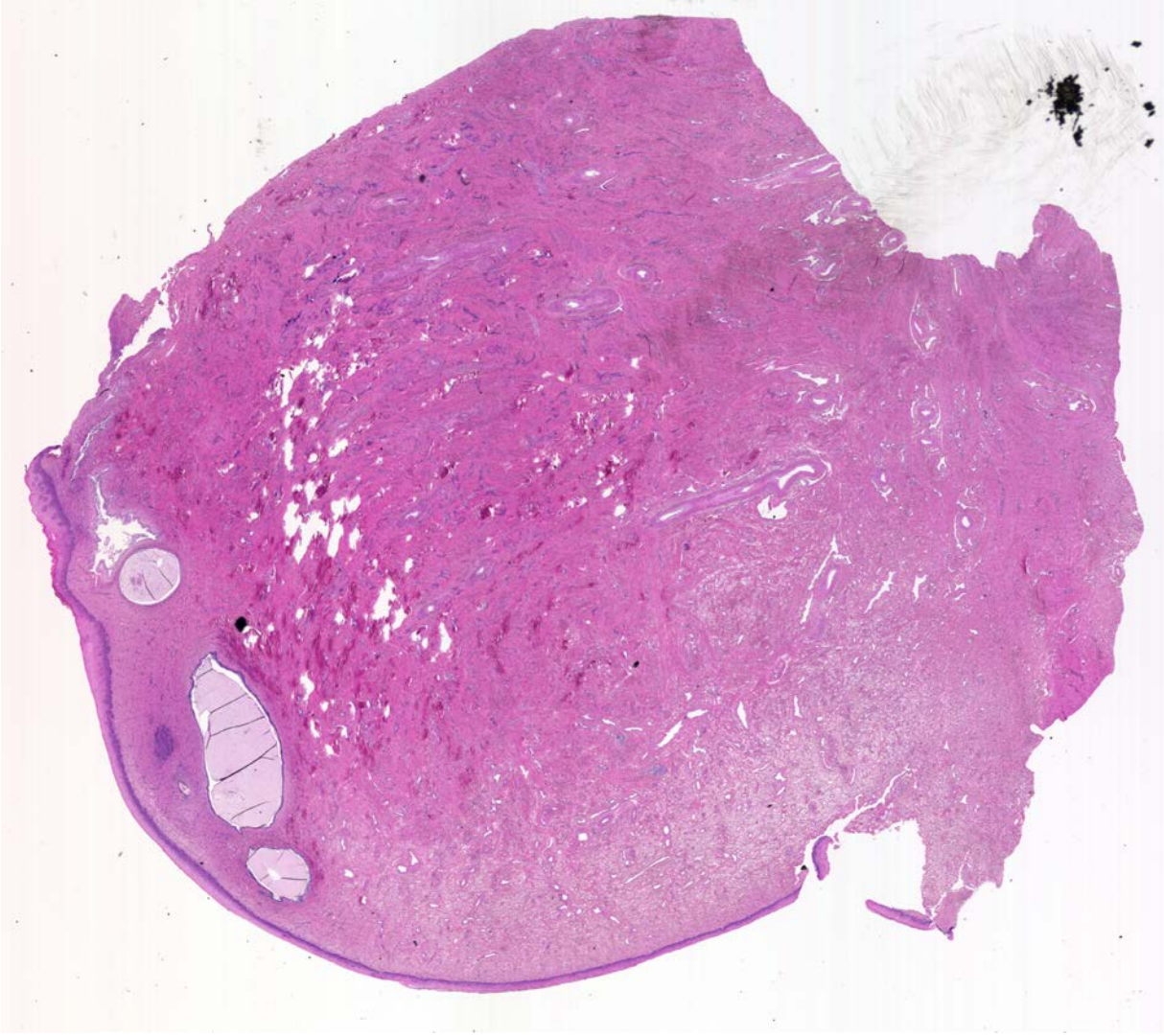


**Figure 2.18. Specimen 12-05-A046a, H&E stain, normal cervical tissue.**

Figure 2.19 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags G and H depict the data collection points on Sample 12-05-A046a; Flag G focuses on the epithelium and the nearby layers while Flag H was the spectra obtained on a deeper tissue layer, the lamina propria.

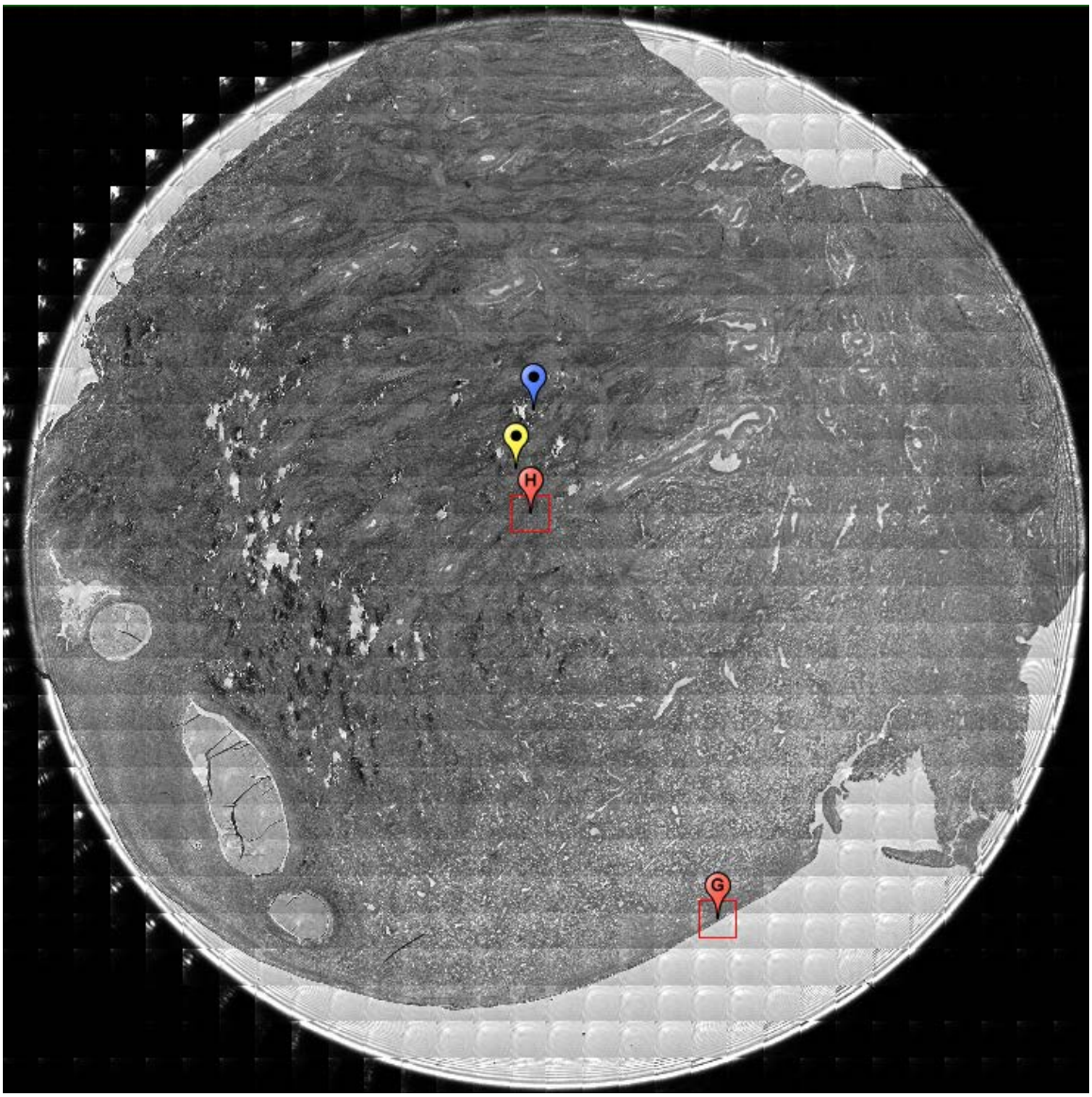



**Figure 2.19. Specimen 12-05-A046a, normal cervical tissue.**

#### 2.3.1.2.1. Sample 12-05-A046a, Point G

A 650 x 650 μm field-of-view image taken of the specimen at Point G (see Figure 2.20) was obtained from Spero's high-magnification objective. The epithelium is visible in this figure, although the single-cell basal layer is not as

evident. While the epithelium on this sample is thin, the basal and parabasal layer and the flat squamous cells in the stratified squamous epithelium are still visible, especially near the birth canal.

Regions of interest were defined, which is where the spectra were actually taken on the sample. Region 1 is identifying an area of interest in the connective tissue of the basement membrane or the lamina propria. Regions 2, 3, 4, 5, and 11 are in the basal and parabasal layer. Regions 6, 7, 8, 9, and 10 are defining the spectral collection points that have the cuboidal and flat squamous cells indicative of the stratified squamous epithelial layer.

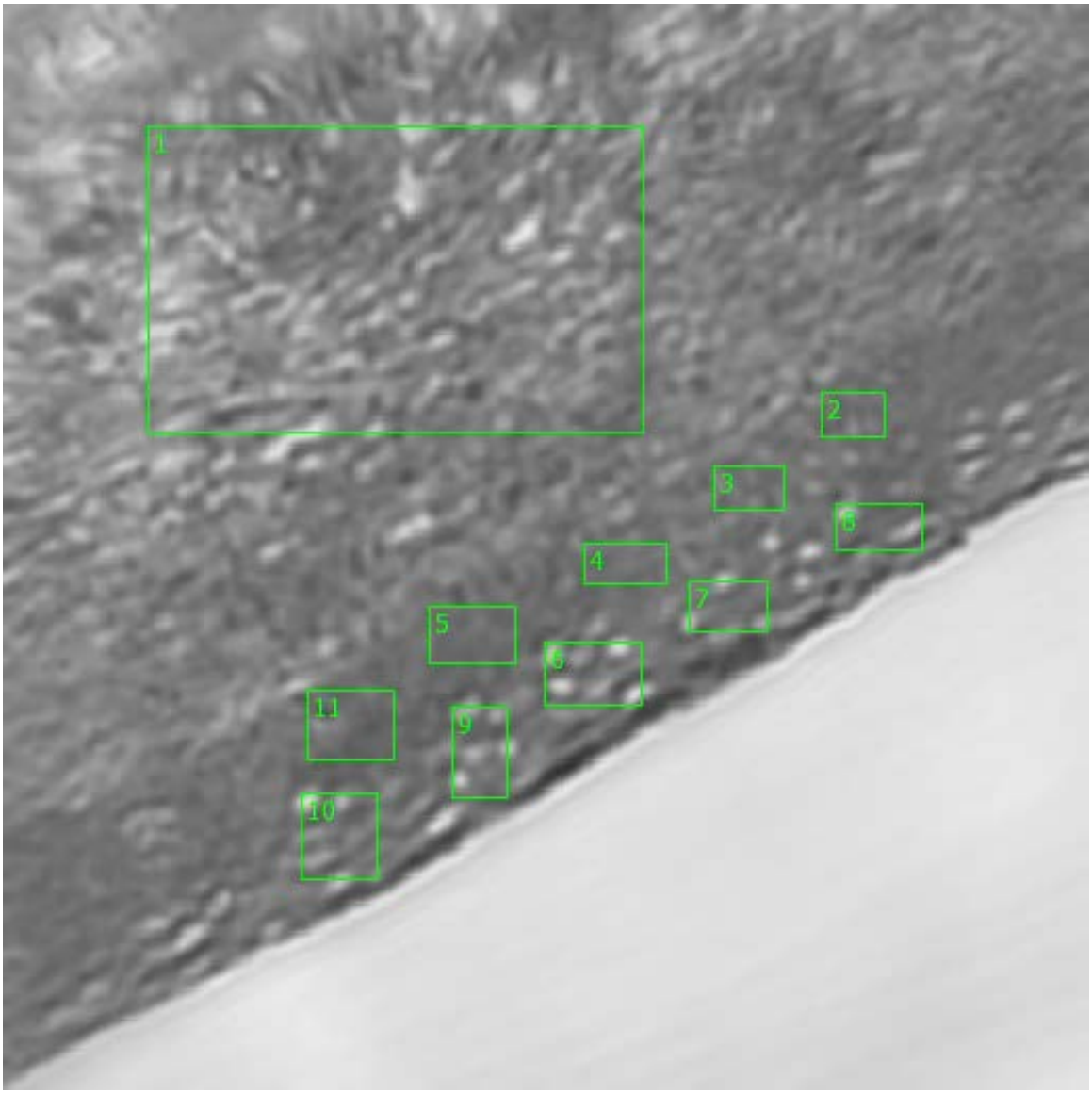



**Figure 2.20. Specimen 12-05-A046a, Point G Regions of Interest**
**Region 1: basement membrane, lamina propria**
**Regions 2, 3, 4, 5, 11: basal layer, parabasal layer**
**Regions 6, 7, 9: columnar epithelium, superficial layer**
**Regions 8, 10: stratified squamous epithelium, superficial layer**

The spectra from the eleven defined regions of interest on Point G are depicted in Figure 2.21.

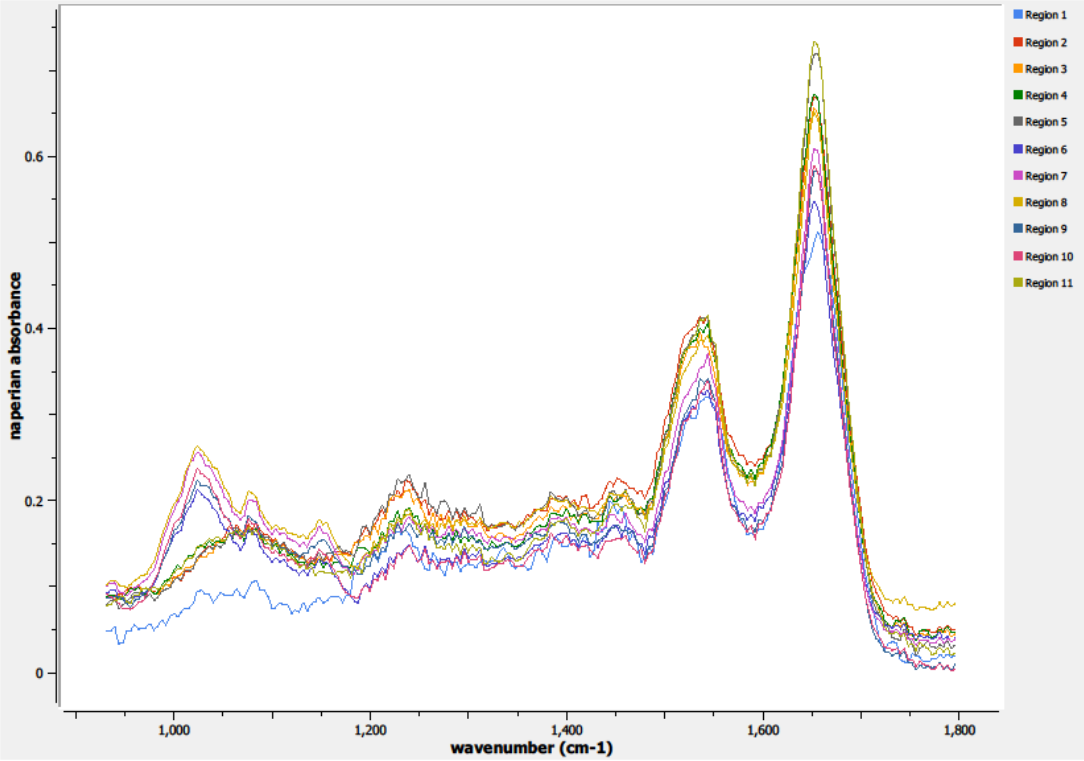



**Figure 2.21. Specimen 12-05-A046a, Point G Spectra**

#### 2.3.1.2.2. Sample 12-05-A046a, Point H

The high-magnification field-of-view image taken at Point H on Sample 12-05-A046a can be seen in Figure 2.22. Point H is located on the interior of the sample, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at this point on the sample.

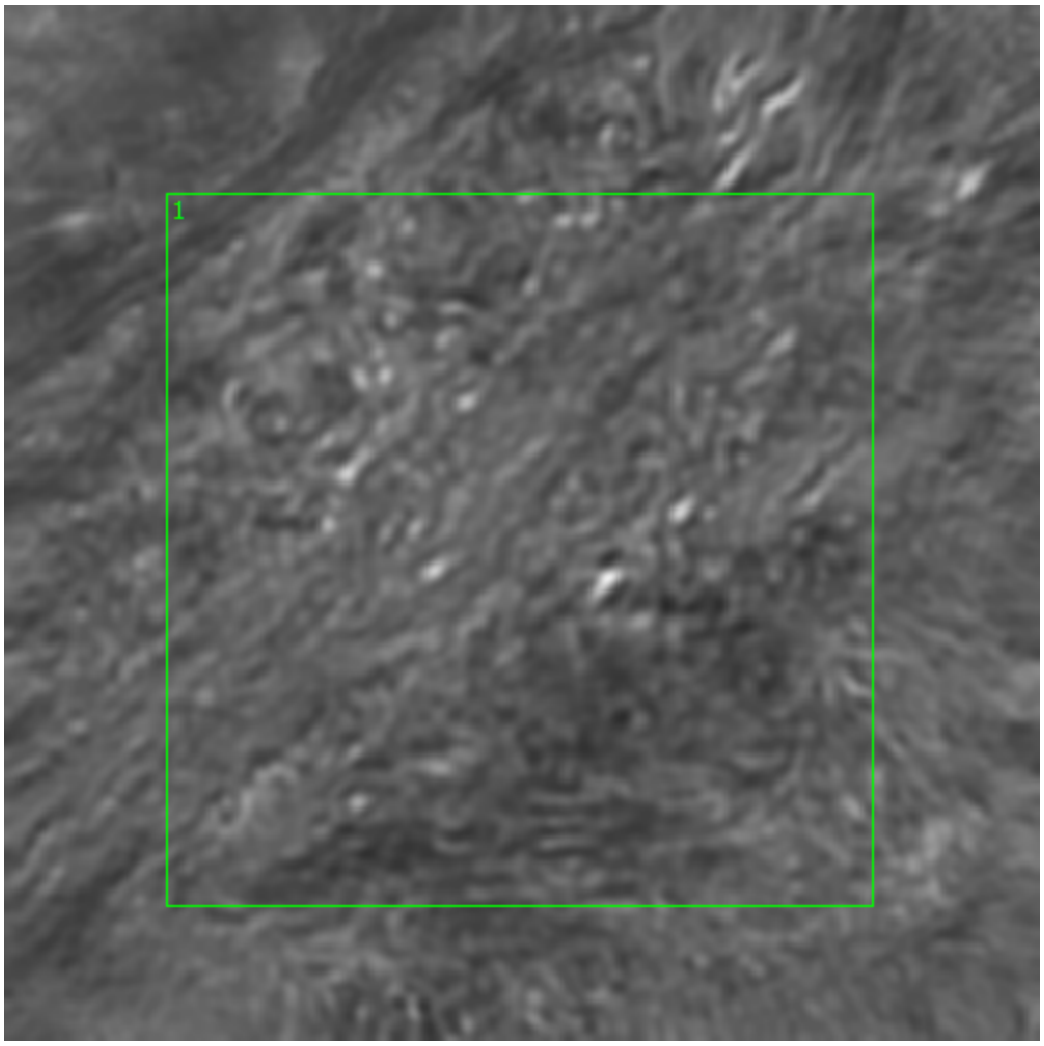



**Figure 2.22. Specimen 12-05-A046a, Point H Regions of Interest**
**Region 1: lamina propria**

The spectrum from this defined region of interest on Point H is depicted in Figure 2.23.

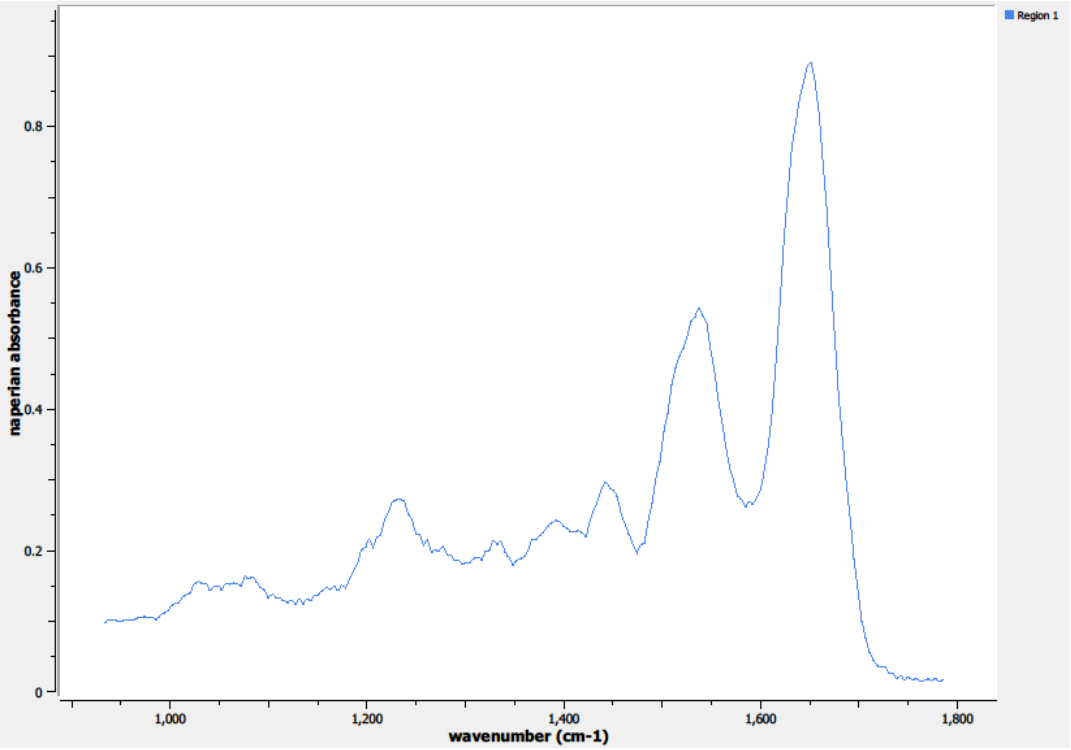



**Figure 2.23. Specimen 12-05-A046a, Point H Spectra**

### 2.3.1.3. Specimen 12-05-A093a

Specimen 12-05-A093a is from a 54-year-old white female. The specimen is a 1.8 gram endocervical sample. Figure 2.24 is an image of a sample of the specimen stained with H&E on a glass slide; a clear epithelium edge is apparent on the tissue sample, as depicted by the dark purple line along the edge of the tissue.

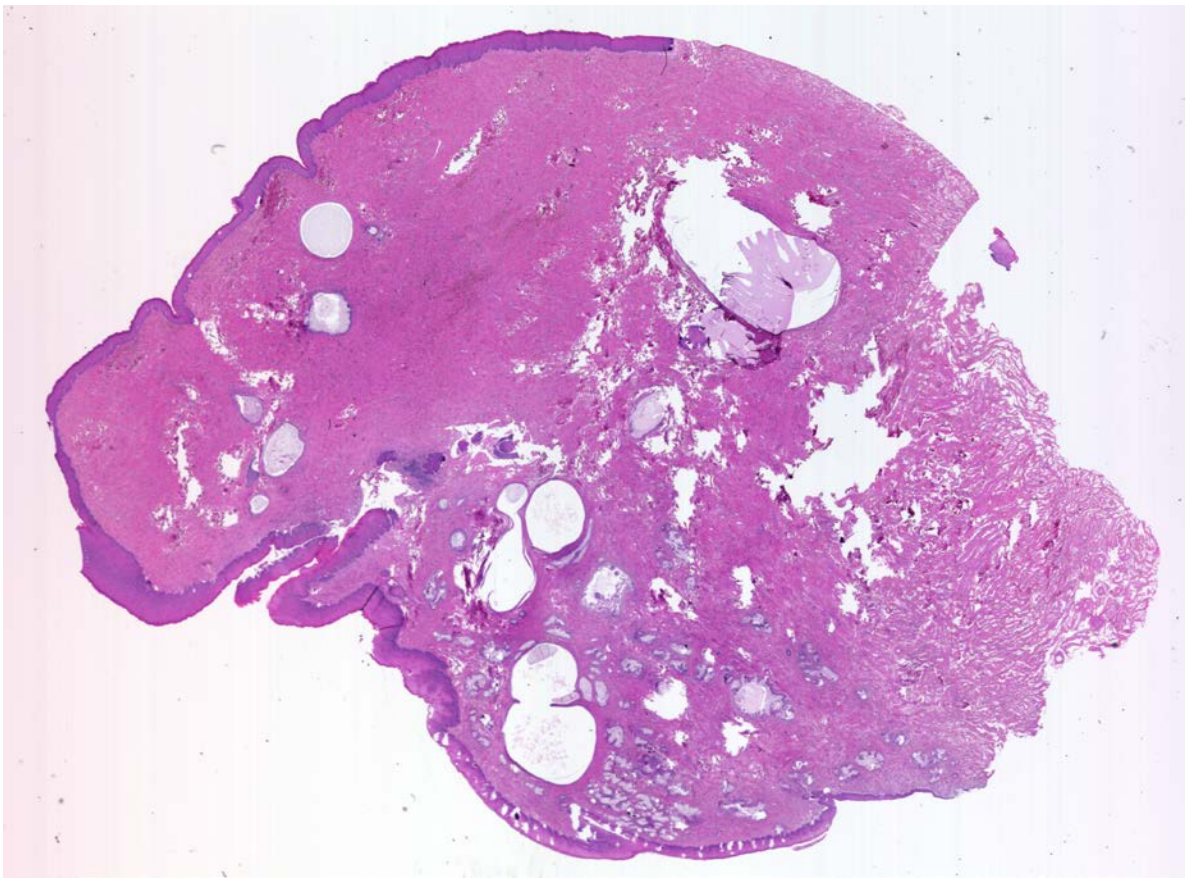


**Figure 2.24. Specimen 12-05-A093a, H&E stain, normal cervical tissue.**

Figure 2.25 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags O, P, Q, and R depict the data collection points on Sample 12-05-A093a; Flags O, P, and Q focus on the epithelium and the nearby layers while Flag R was the spectra obtained on a deeper tissue layer, the lamina propria.

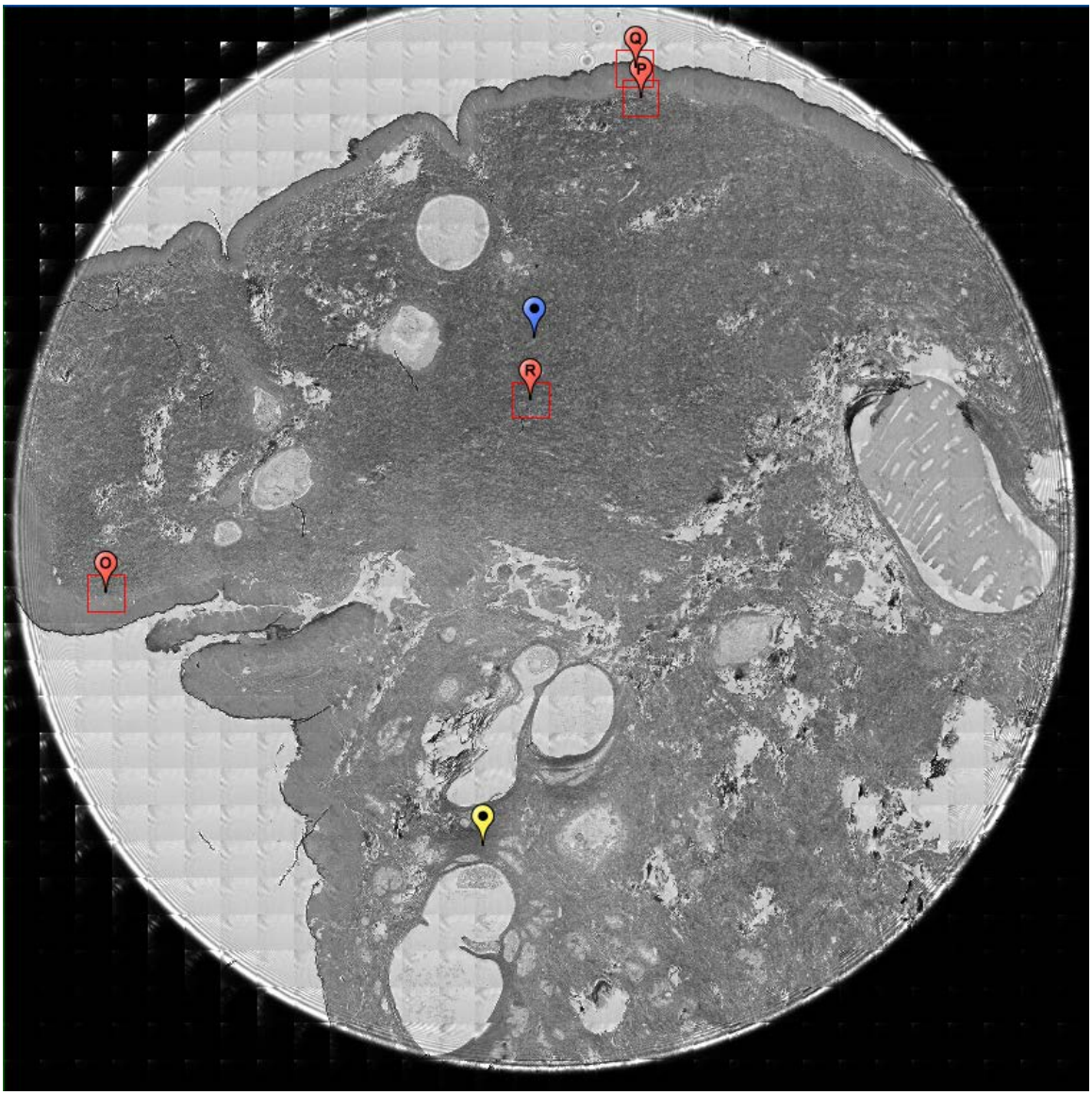



**Figure 2.25. Specimen 12-05-A093a, normal cervical tissue.**

#### 2.3.1.3.1. Specimen 12-05-A093a, Point O

A 650 x 650 µm field-of-view image taken of the specimen at Point O (see Figure 2.26) was obtained from Spero's high-magnification objective. The basal cellular line separating the epithelium from the basement membrane is fainter than

previous images; the “fingers” of the basement membrane can still be seen invading the epithelium. Region 1 is taken in the basement membrane or the lamina propria; Regions 2 and 3 are taken just beyond the single-cell basal layer, stretching into the parabasal layer.

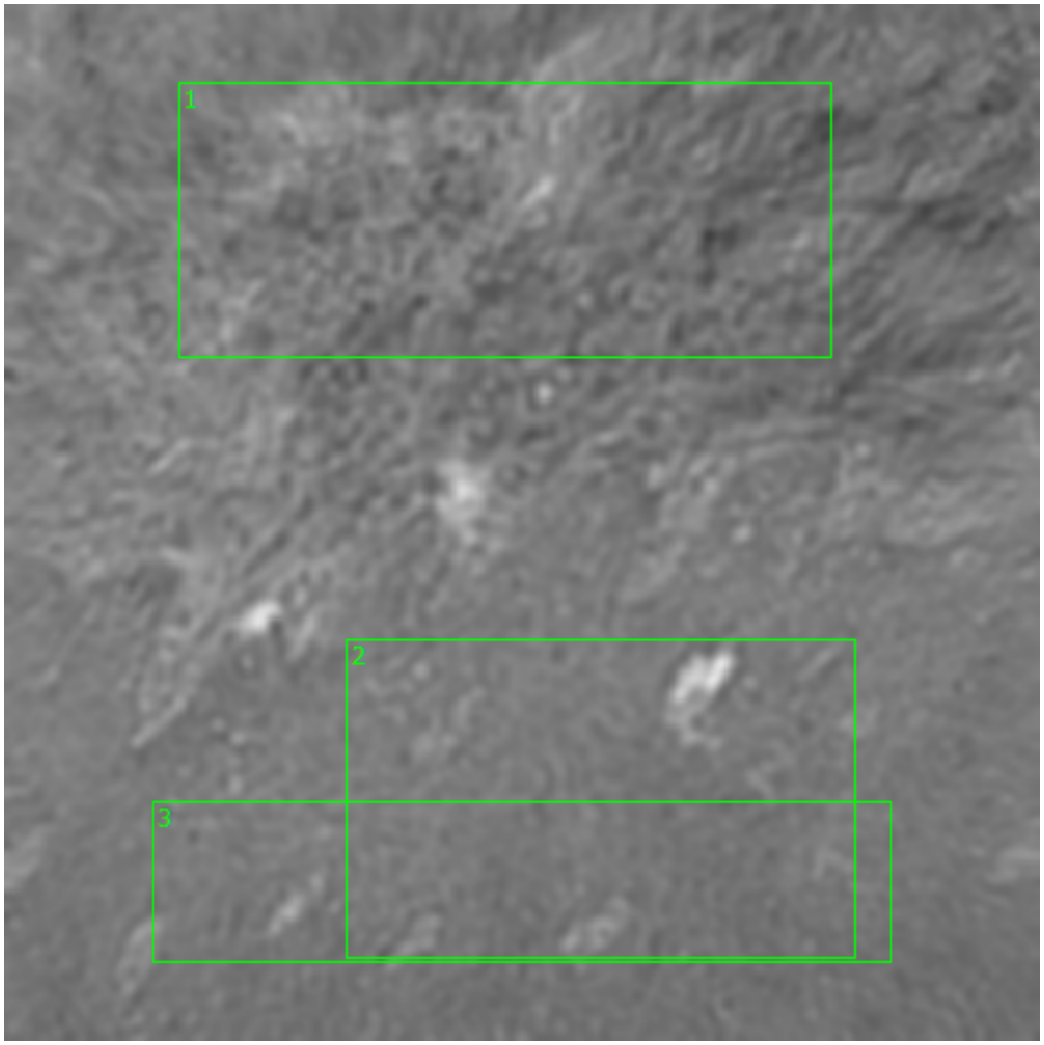



**Figure 2.26. Specimen 12-05-A093a, Point O Regions of Interest**
**Region 1: basement membrane, lamina propria**
**Regions 2, 3: basal layer, parabasal layer**

The spectra from the three defined regions of interest on Point O are depicted in Figure 2.27.

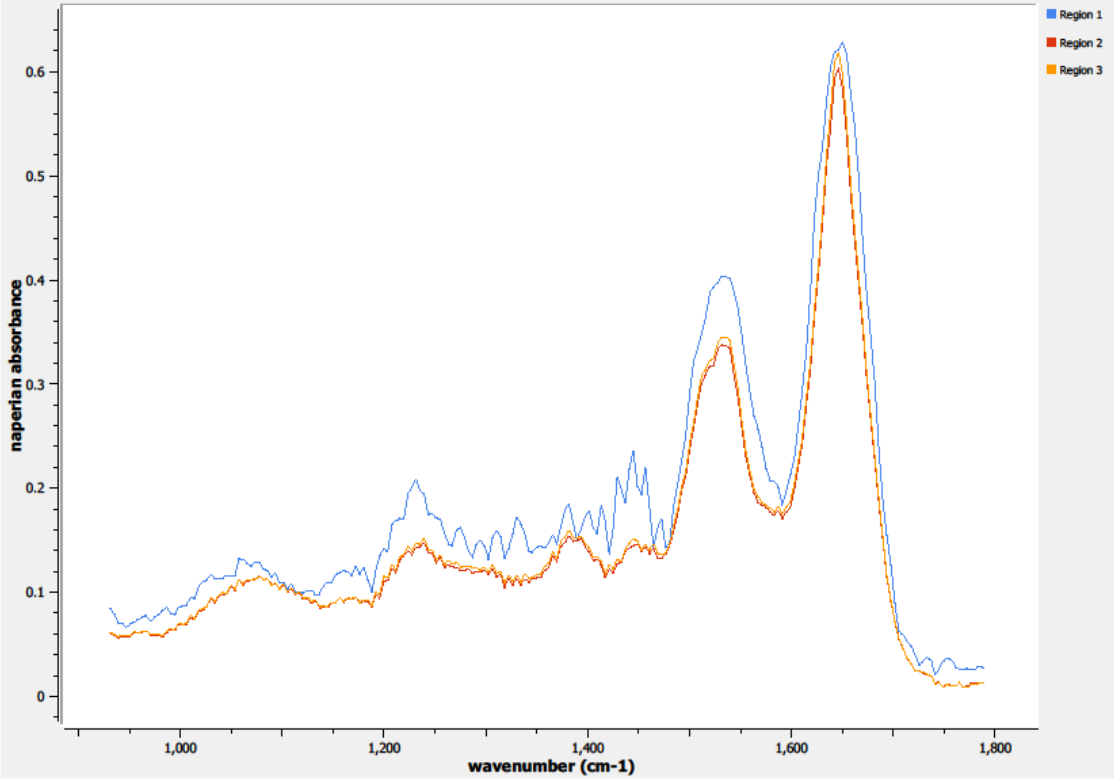



**Figure 2.27. Specimen 12-05-A093a, Point O Spectra**

#### 2.3.1.3.2. Sample 12-05-A093a, Point P

The high-magnification 650 x 650 µm field-of-view image was also taken at Point P (see Figure 2.28) on Sample 12-05-A093a. The basal cellular layer is again relatively faint, although the tissue type is clearly distinguishable between the epithelium and the basement membrane.

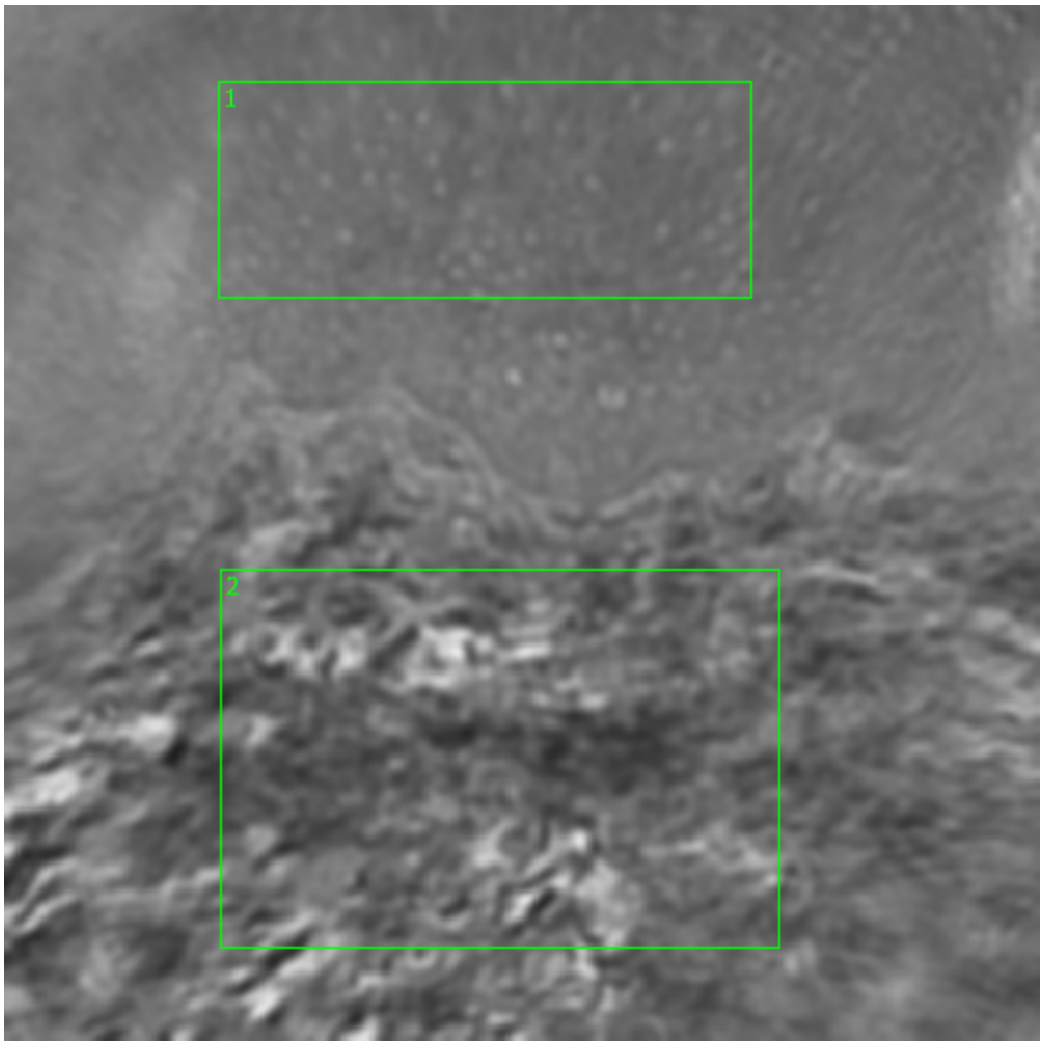



**Figure 2.28. Specimen 12-05-A093a, Point P Regions of Interest**
**Region 1: parabasal layer**
**Region 2: basement membrane, lamina propria**

Region 1 is located in the epithelium of Point P. Cuboidal cells are evident in this layer, which encompasses the parabasal region of the epithelium. Stratified squamous cells start appearing at the edge of the image, but per the recommendations of the microscope's manufacture, spectral data was not collected near the edge of the image to avoid optical artifacts in Spero.

Region 2 is identifying regions of interest in the basement membrane and the lamina propria. The connective tissue layer is visually identifiable from the cellular structures in the nearby epithelium.

The spectra from the two defined regions of interest on Point P are depicted in Figure 2.29.

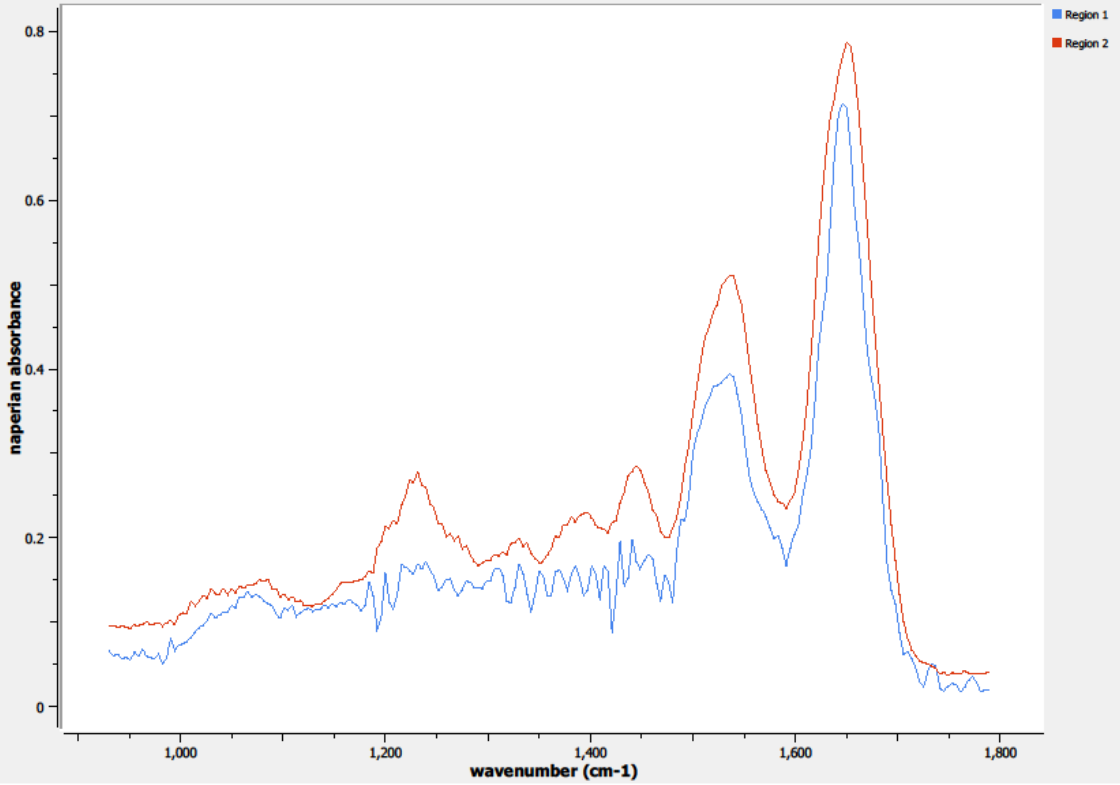



**Figure 2.29. Specimen 12-05-A093a, Point P Spectra**

#### 2.3.1.3.3. Sample 12-05-A093a, Point Q

The high-magnification field-of-view image taken at Point Q on Sample 12-05-A093a can be seen in Figure 2.30. At this point in the tissue sample, the

epithelium is quite thick, which is why Point Q is located directly adjacent to Point P to fully cover the basement membrane and epithelium.

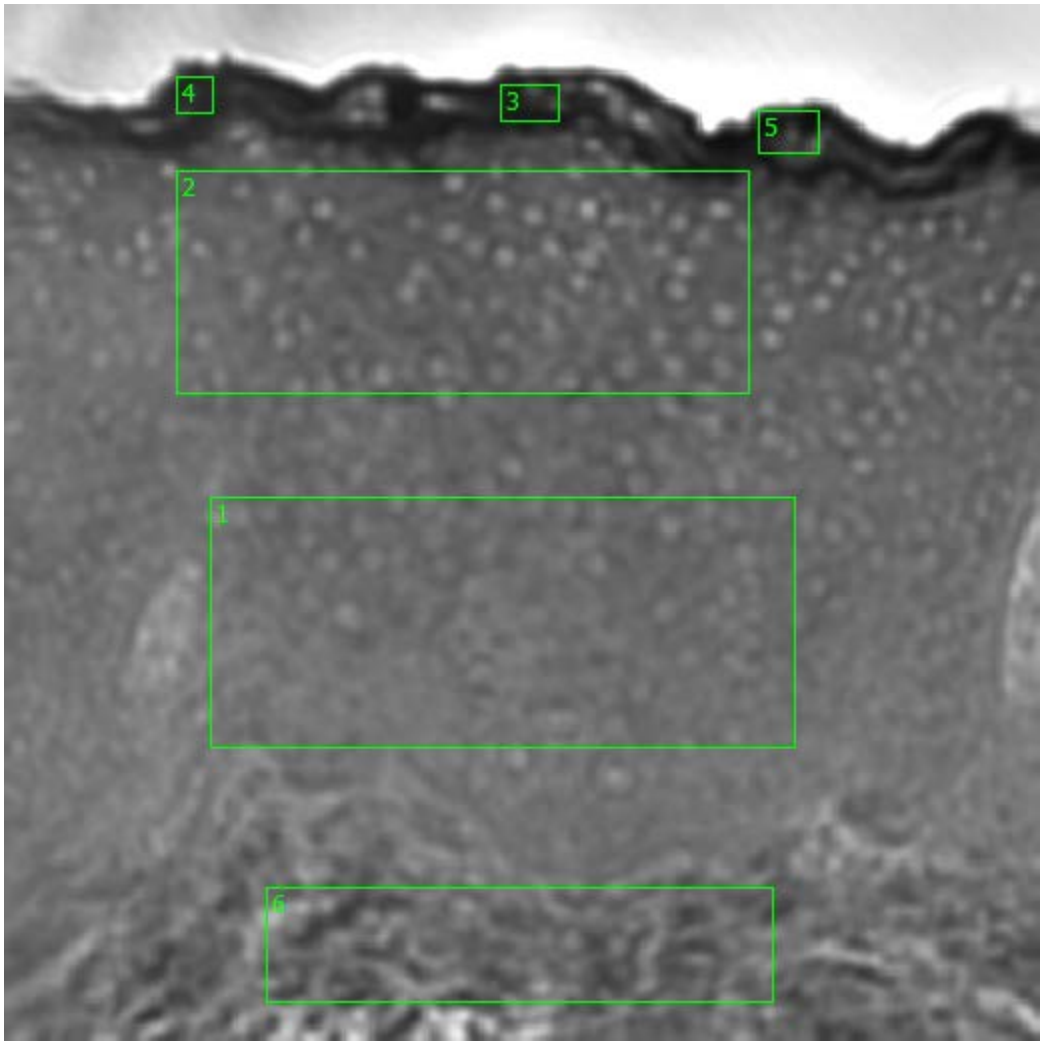



**Figure 2.30. Specimen 12-05-A093a, Point Q Regions of Interest**
**Region 1: parabasal layer, columnar epithelium**
**Region 2: columnar epithelium**
**Regions 3, 4, 5: stratified squamous epithelium, epithelial surface**
**Region 6: basement membrane**

The basal cellular layer is faint on Point P, but the tissue type is distinguishable between the epithelium and the cellular structure of the basement membrane. Region 1 is located on parabasal region of the epithelium of Point Q. Region 2 is not as deep in the epithelium as Region 1; cuboidal cells are manifest in this region of interest.

Regions 3, 4, and 5 are located in the epithelial surface of Point Q on Sample 12-05-A093a, adjacent to the lower portion of the uterus. High absorption is expected in this layer. A small quantity of stratified squamous cells starts appearing at the outer edge of the epithelium.

Region 6 is identifying a region of interest in the basement membrane. The connective tissue layer is visually identifiable from the cellular structures in Point Q's adjoining epithelium.

The spectra from the six defined regions of interest on Point Q are depicted in Figure 2.31.

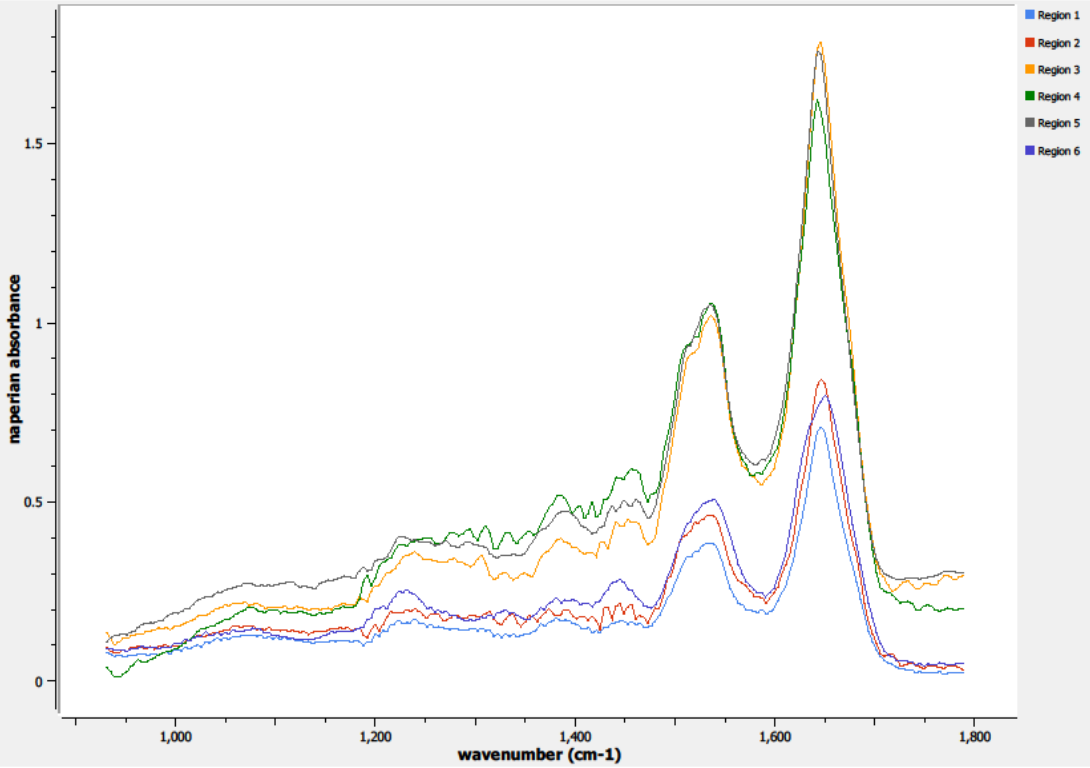



**Figure 2.31. Specimen 12-05-A093a, Point Q Spectra**

#### 2.3.1.3.4. Sample 12-05-A093a, Point R

The high-magnification field-of-view image taken at Point R on Sample 12-05-A093a can be seen in Figure 2.32. Point R is located on the interior of the sample, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at this point on the sample.

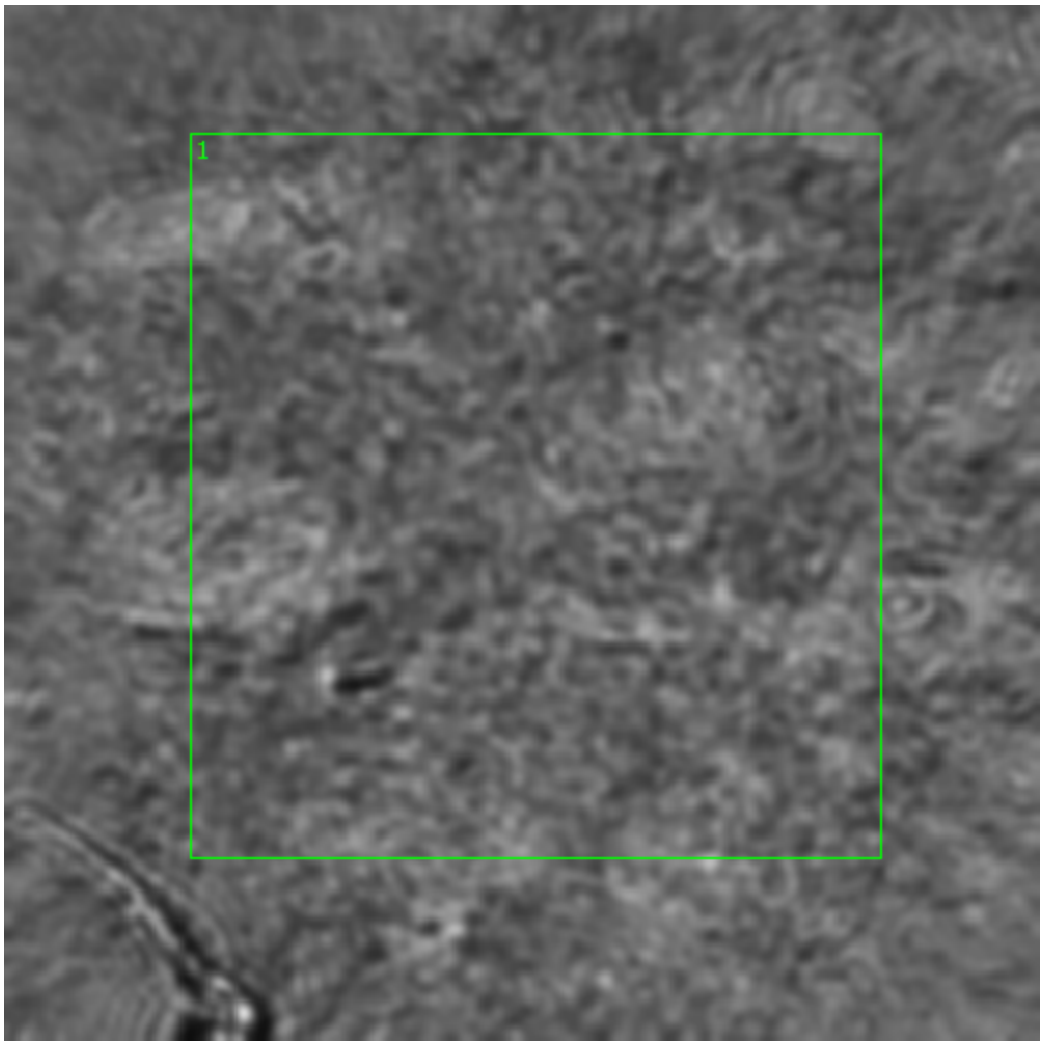



**Figure 2.32. Point R Regions of Interest**
**Region 1: lamina propria**

The spectrum from this defined region of interest on Point R is depicted in Figure 2.33.

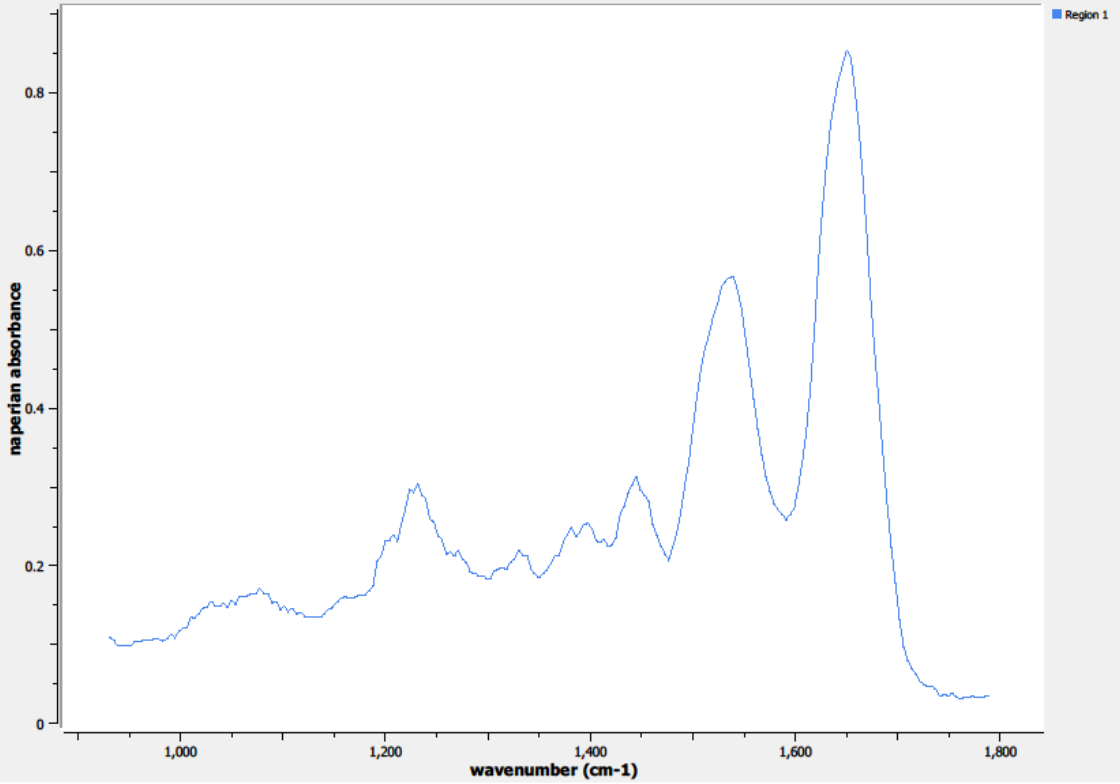



**Figure 2.33. Point R Spectra**

#### 2.3.1.4. Specimen 12-05-A072a

Specimen 12-05-A072a is from a 55-year-old Hispanic female. The specimen is a 0.42 gram sample in the junction between the ectocervix and the endocervix. This specimen is unique in my normal cervical sample lot in the following two ways: it is the only Hispanic specimen, and it is my only ecto/endocervical junction. Figure 2.34 is an image of a sample of the specimen stained with H&E on a glass slide.

**Figure 2.34. Specimen 12-05-A072a, H&E stain, normal cervical tissue.**

A small portion of the H&E-stained sample has an epithelium layer in the lower left-hand corner of Figure 2.34, which can be seen by the darker purple line near the edge of the tissue sample. The right side of the tissue sample has a portion that is a darker purple, indicating a higher tissue density. This is often indicative of a tumor growth or the start of an abnormal growth; regardless, that section of the sample appears abnormal, so spectra were not collected from that portion of the tissue.

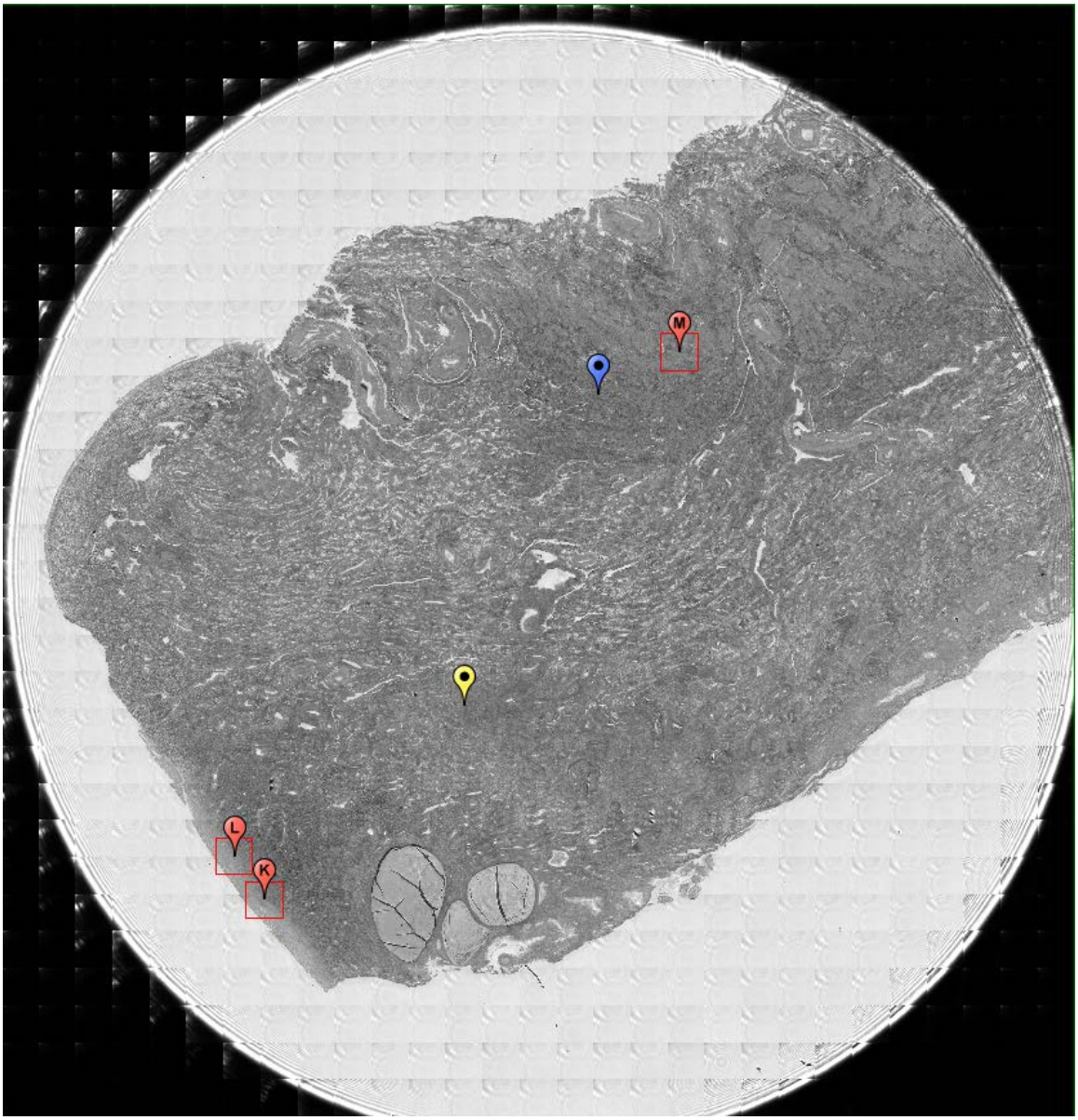



**Figure 2.35. Specimen 12-05-A072a, normal cervical tissue.**

Figure 2.35 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags K, L, and M depict the data collection points on Sample 12-05-A072a; Flags K and L focus on the

epithelium and the nearby layers while Flag M was the spectra obtained on a deeper tissue layer, the lamina propria.

#### 2.3.1.4.1. Specimen 12-05-A072a, Point K

A 650 x 650 μm field-of-view image taken of the specimen at Point K on Sample 12-05-A072a (see Figure 2.36) was obtained from Spero's high-magnification objective. The basal cellular layer is faint on Point K, but the tissue type is distinguishable between the epithelium and the cellular structure of the basement membrane.

Regions 2, 3, 4, 5, and 6 are defining areas of interest just beyond the single-cell basal layer, stretching into the parabasal layer. Regions 7, 8, and 9 are taken in the parabasal layer, stretching into the intermediate layer of the epithelium. Regions 10 and 11 have more defined columnar and squamous cells on the superficial epithelial layer. Finally, Region 12 is identifying a region of interest in the connective tissue of the basement membrane of the cervix on Point K.

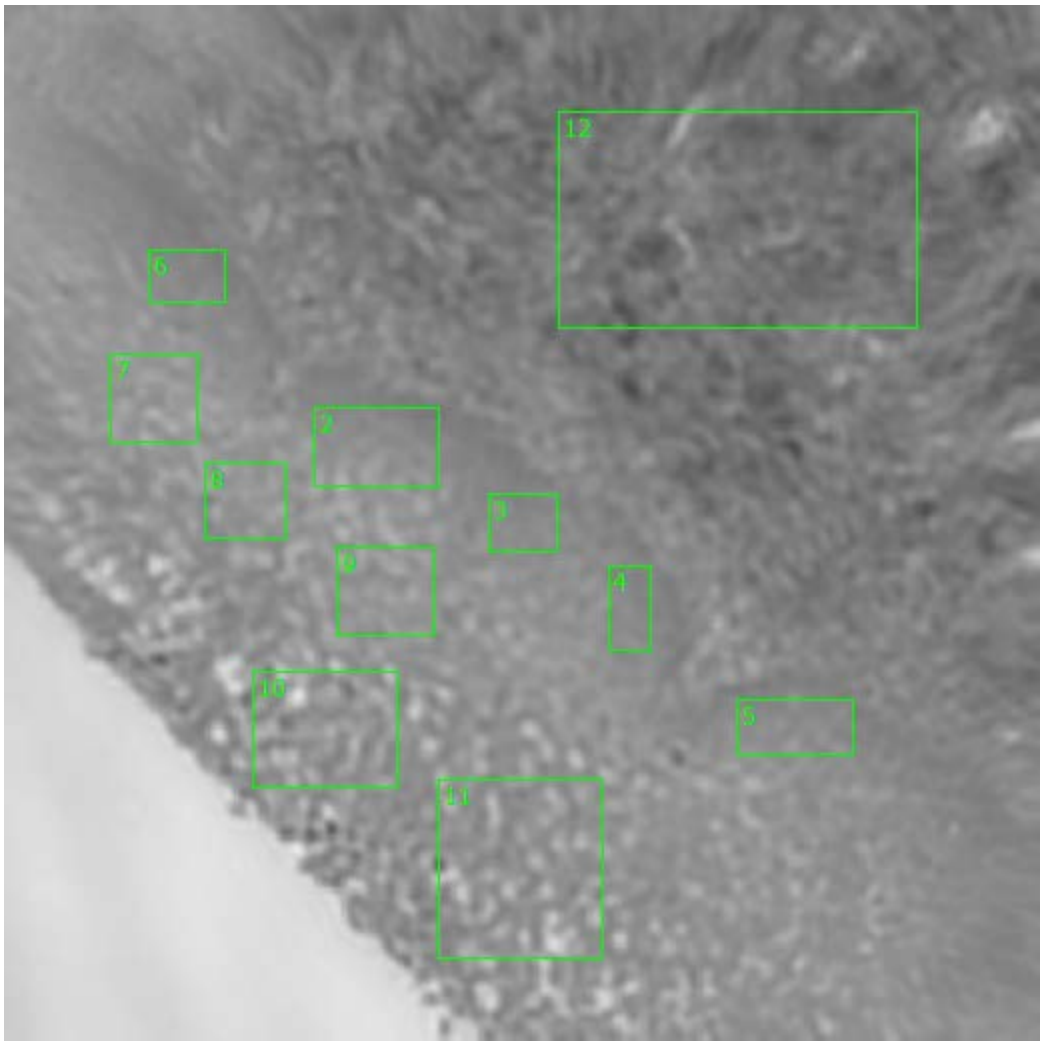



**Figure 2.36. Specimen 12-05-A072a, Point K Regions of Interest**
**Regions 2, 3, 4, 5, 6: basal layer, parabasal layer**
**Regions 7, 8, 9: columnar epithelium, intermediate layer**
**Regions 10, 11: columnar epithelium, superficial layer**
**Region 12: basement membrane, lamina propria**

The spectra from the eleven defined regions of interest on Point K are depicted in Figure 2.37.

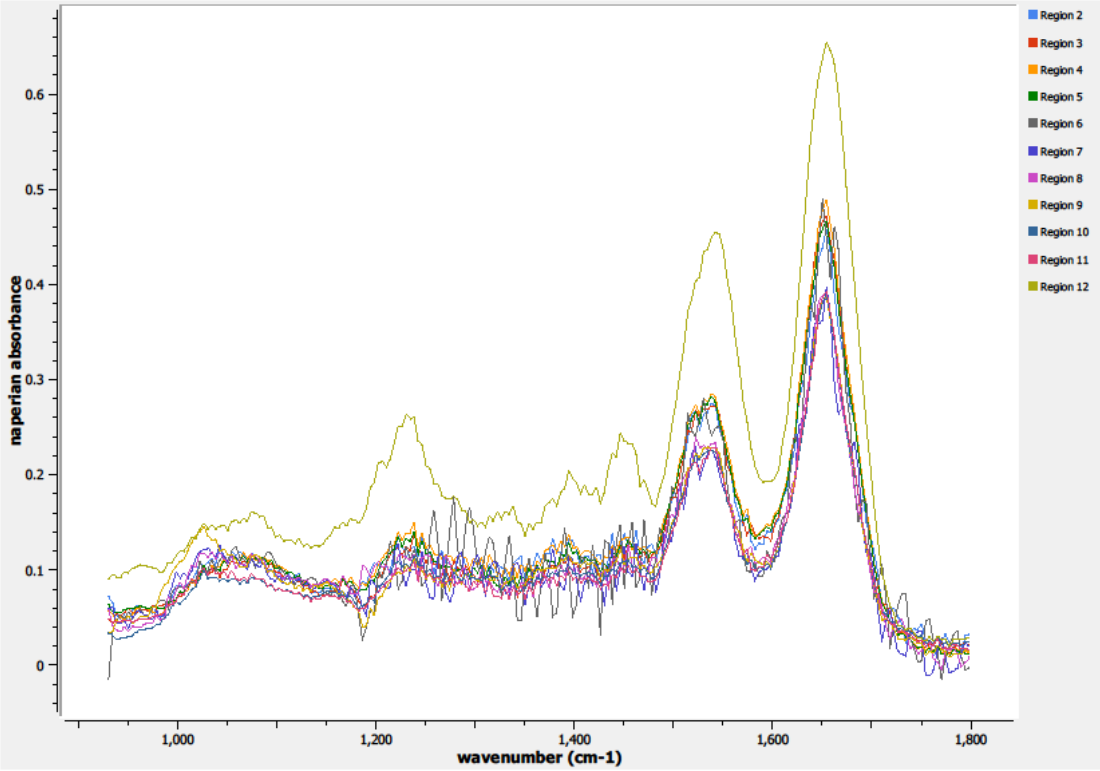



**Figure 2.37. Specimen 12-05-A072a, Point K Spectra**

#### 2.3.1.4.2. Specimen 12-05-A072a, Point L

The high-magnification 650 x 650 µm field-of-view image was also taken at Point L (see Figure 2.38) on Sample 12-05-A072a. The basal cellular layer is again relatively faint on Point L, although the tissue type is clearly distinguishable between the epithelium and the connective tissue cellular structure of the basement membrane.

Region 1 is identifying a region of interest in the connective tissue of the basement membrane and lamina propria of the cervix on Point L. Regions 2, 3, and 4 are defining areas of interest just beyond the single-cell basal layer, stretching into

the parabasal layer. Regions 5 and 6 have more defined cuboidal and squamous cells on the superficial epithelial layer on Point L.

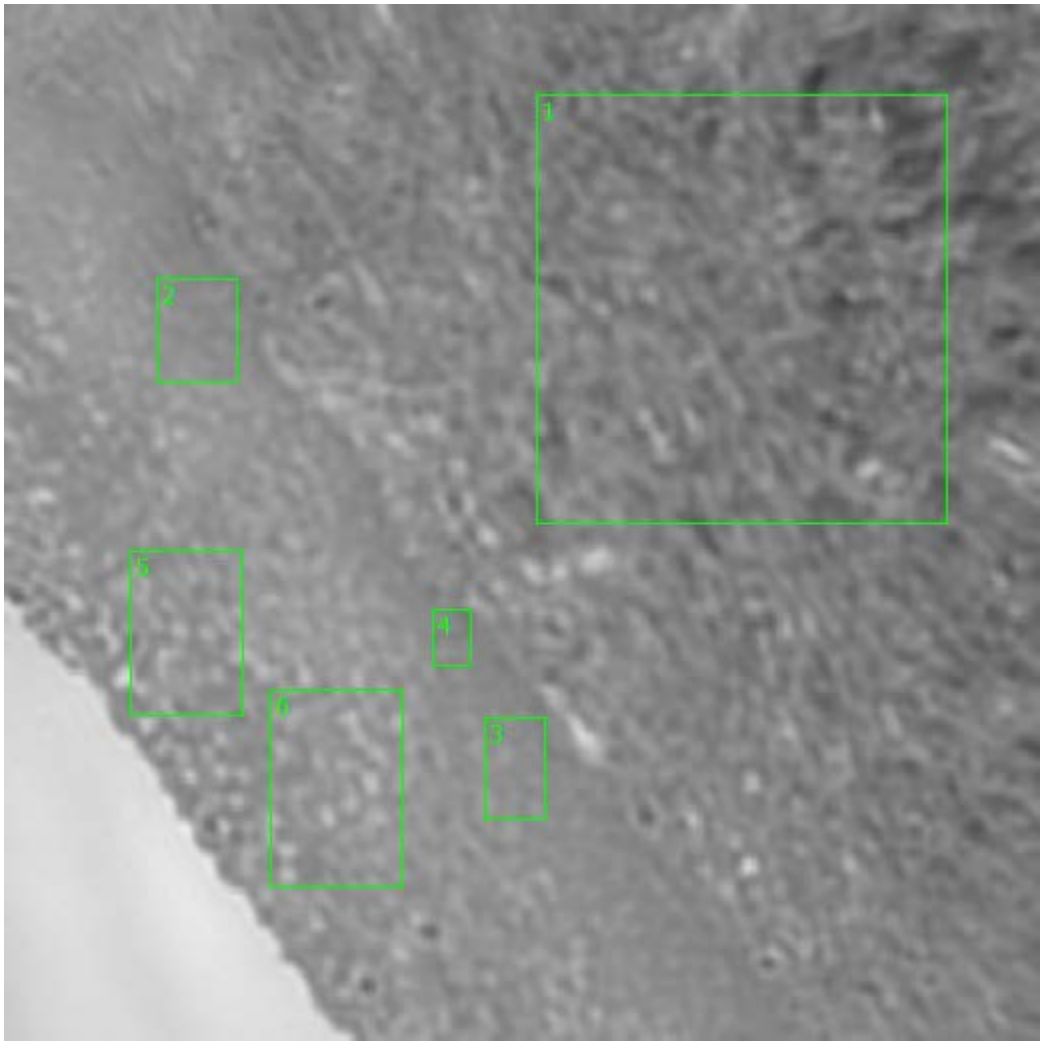



**Figure 2.38. Specimen 12-05-A072a, Point L Regions of Interest**
**Region 1: basement membrane, lamina propria**
**Regions 2, 3, 4: basal layer, parabasal layer**
**Regions 5, 6: columnar epithelium, superficial layer**

The spectra from the six defined regions of interest on Point L are depicted in Figure 2.39.

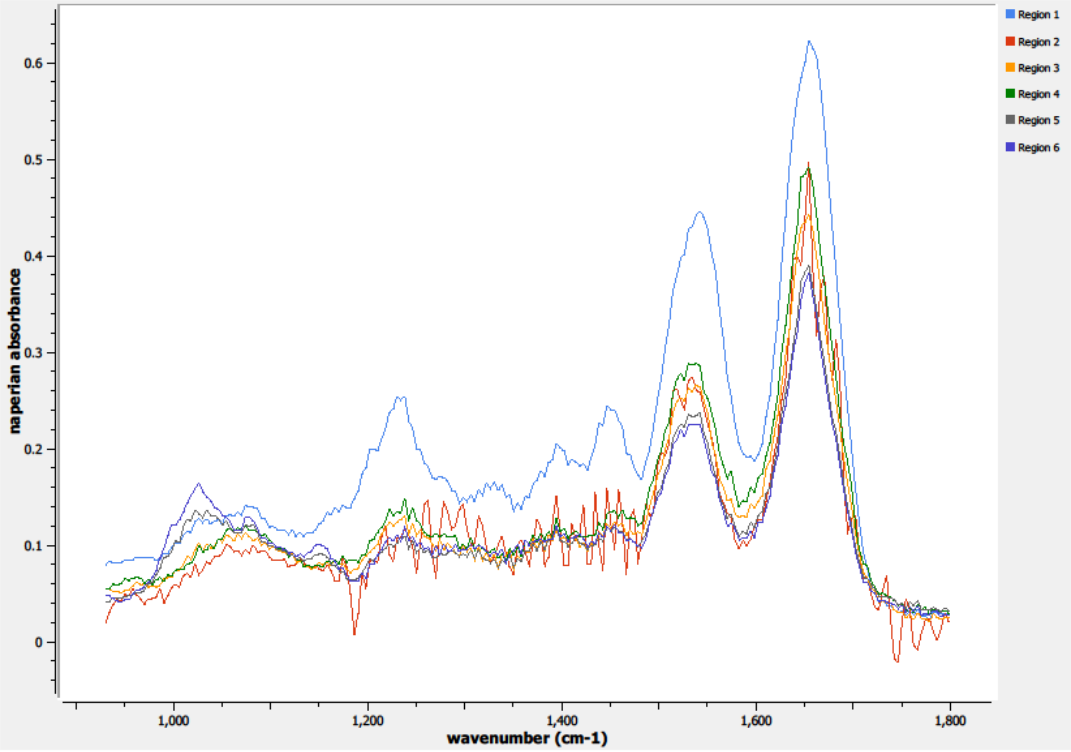



**Figure 2.39. Specimen 12-05-A072a, Point L Spectra**

#### 2.3.1.4.3. Specimen 12-05-A072a, Point M

The high-magnification field-of-view image taken at Point M on Sample 12-05-A072a can be seen in Figure 2.40. Point M is located on the interior of the sample, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at this point on the sample.

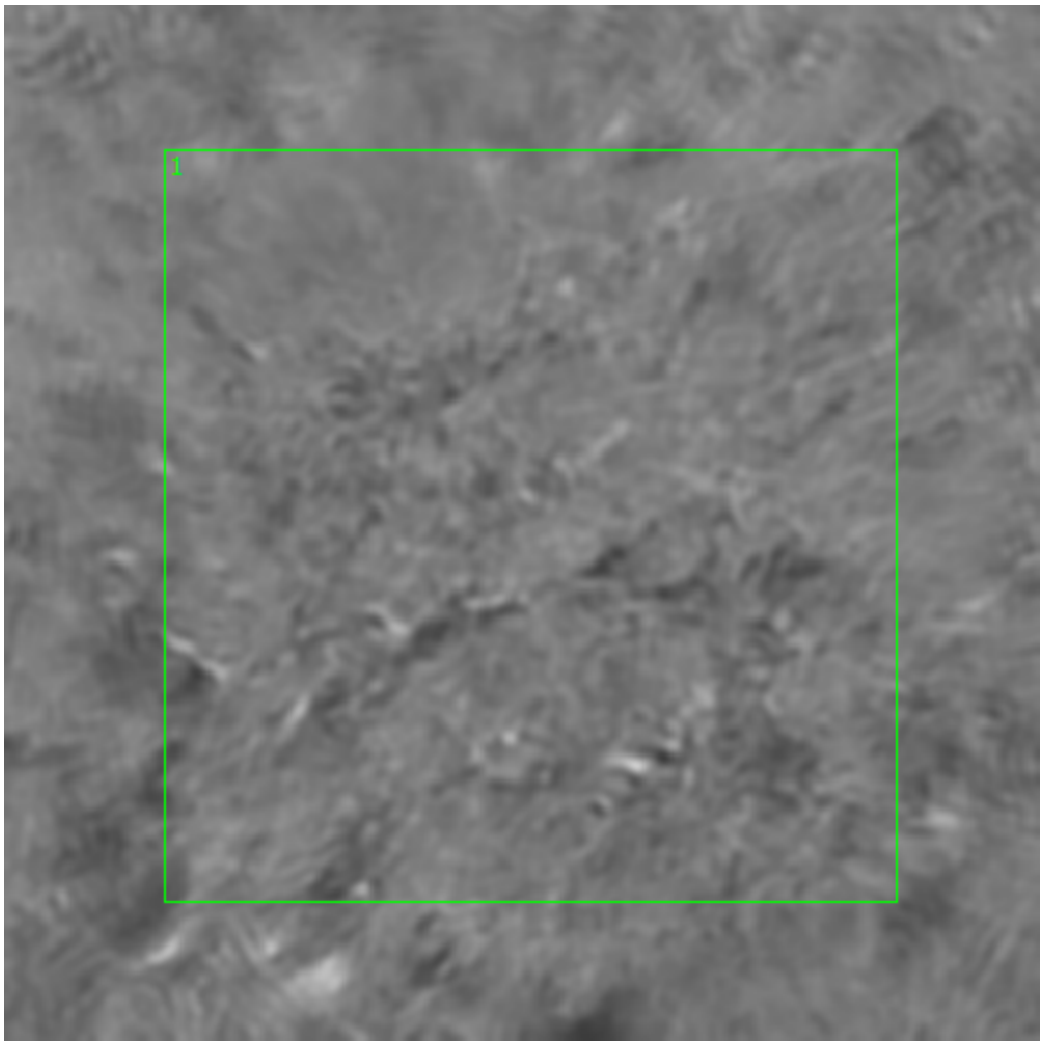



**Figure 2.40. Specimen 12-05-A072a, Point M Regions of Interest**
**Region 1: lamina propria**

The spectrum from this defined region of interest on Point M is depicted in Figure 2.41.

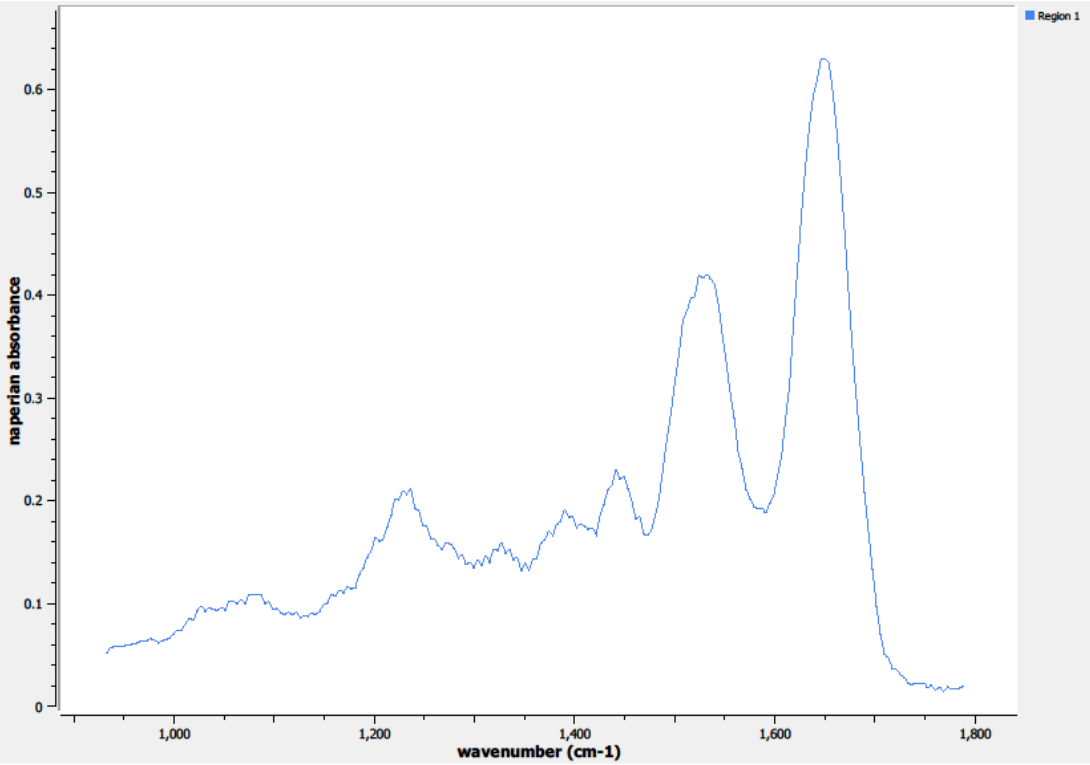



**Figure 2.41. Specimen 12-05-A072a, Point M Spectra**

### 2.3.2. Malignant Cervical Data Acquisition

Fourteen malignant cervical specimens were utilized for this experiment. Samples from each specimen were placed on glass slides and CaF2 windows; the samples on the glass slides were stained with H&E. Data for each sample on the CaF2 windows were collected on Spero and evaluated in this experiment; the locations of the data collection points were first identified on the samples stained with H&E. Each sample had multiple points of data collection; a few key malignant cervical specimen examples are listed in this section. Refer to Appendix B for a

complete listing of malignant cervical specimens, including images from the H&E-stained slides, the images obtained through Spero's high-magnification objective, and the spectra from each data point collected.

#### 2.3.2.1. Specimen 93-02-A251

Specimen 93-02-A251 is a malignant cervical sample from a 22-year-old white female. CHTN did not include a weight of the sample or a cervical location from which the specimen was removed. The quality control diagnostic provided by CHTN personnel identifies the specimen as squamous cell carcinoma of the cervix; the tumor percentage was not provided with Specimen 93-02-A251.

An image of a sample from Specimen 93-02-A251 stained with H&E on a glass slide is shown in Figure 2.42. The dark purple throughout the image is indicative of high cell concentration typically associated with a tumor growth.

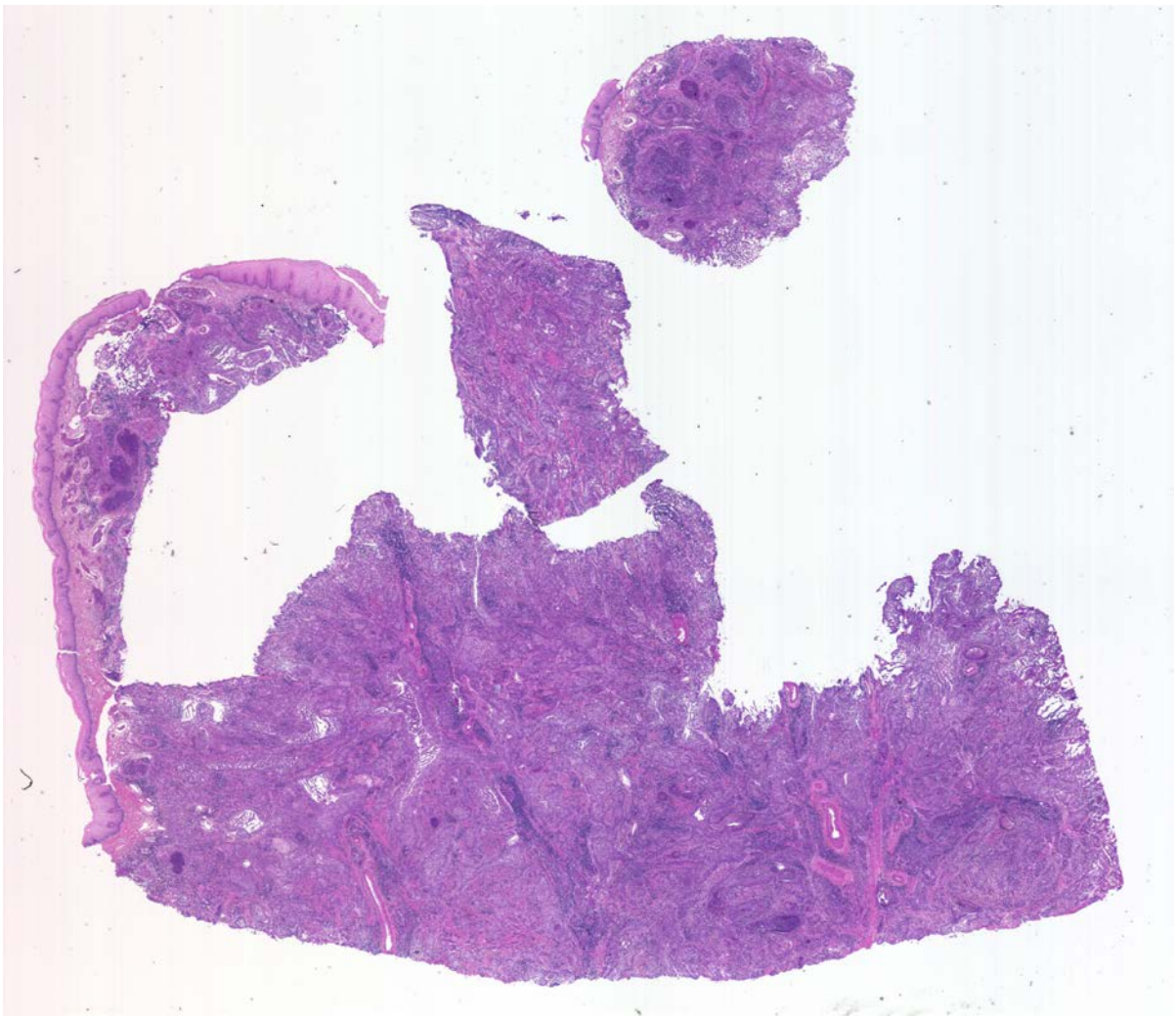


**Figure 2.42. Specimen 93-02-A251, H&E stain, squamous cell carcinoma of the cervix**

Figure 2.43 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags E and F depict the spectral data collection points on Sample 93-02-A251, focusing on areas of higher cell concentration.

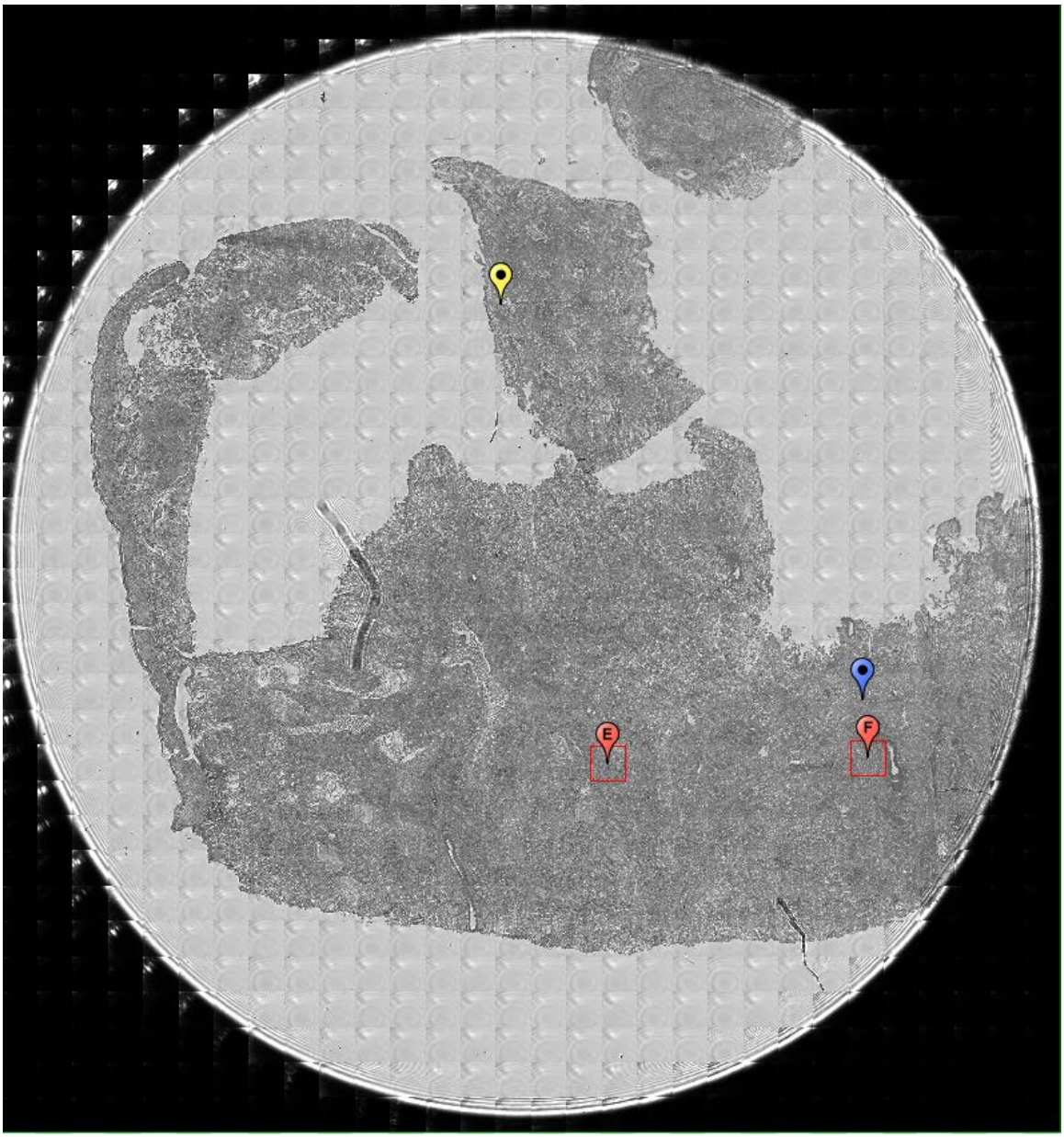



**Figure 2.43. Specimen 93-02-A251, cervical cancer tissue**

#### 2.3.2.1.1. Sample 93-02-A251, Point E

A 650 x 650 μm field-of-view image taken of the specimen at Point E (see Figure 2.44) was obtained from Spero's high-magnification objective. Point E is

located on the interior of Sample 93-02-A251, in the lamina propria layer. Point E's location on the sample was targeting the sample's higher cell concentration areas. Only one region of interest was needed to capture the spectral information at this point on the sample.

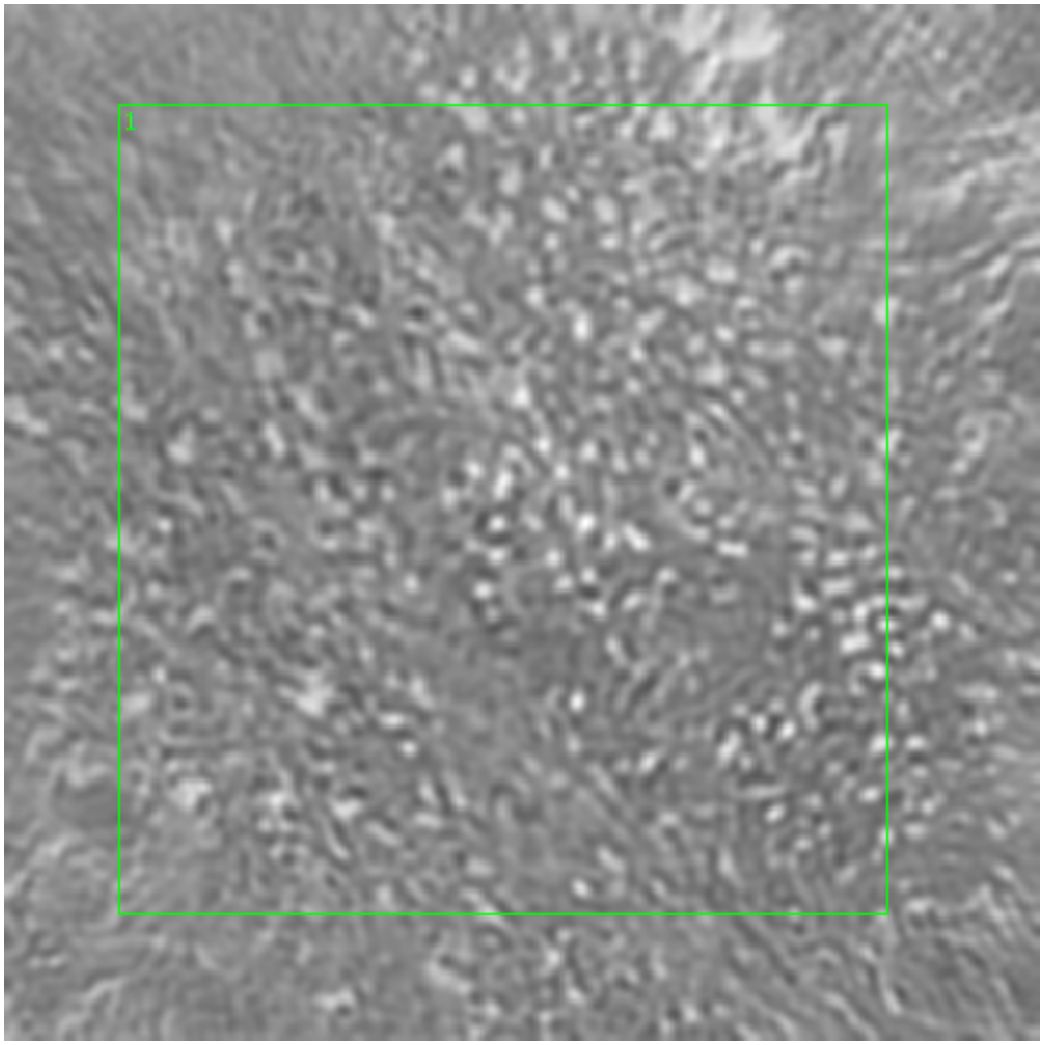



**Figure 2.44. Specimen 93-02-A251, Point E Regions of Interest**
**Region 1: lamina propria, squamous cell carcinoma**

The spectrum from the defined region of interest on Point E is depicted in Figure 2.45.

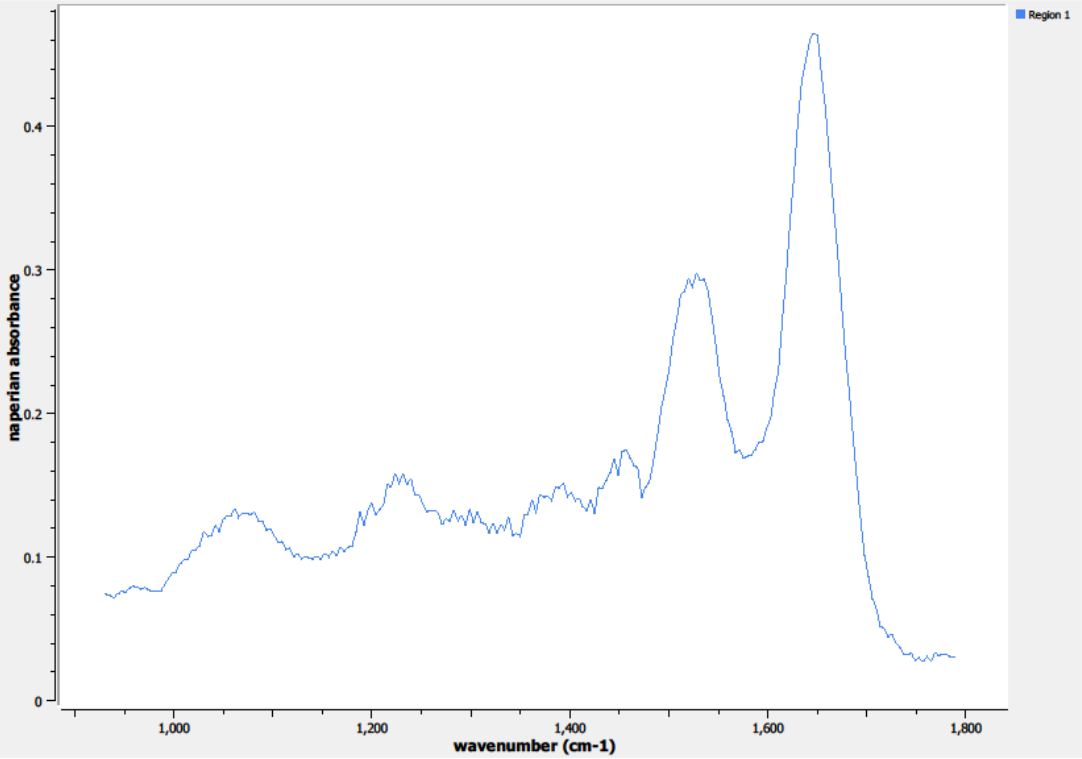



**Figure 2.45. Specimen 93-02-A251, Point E Spectra**

#### 2.3.2.1.2. Sample 93-02-A251, Point F

The high-magnification field-of-view image taken at Point F on Sample 93-02-A251 can be seen in Figure 2.46. Point F is located on the interior of the sample, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at this point on the sample.

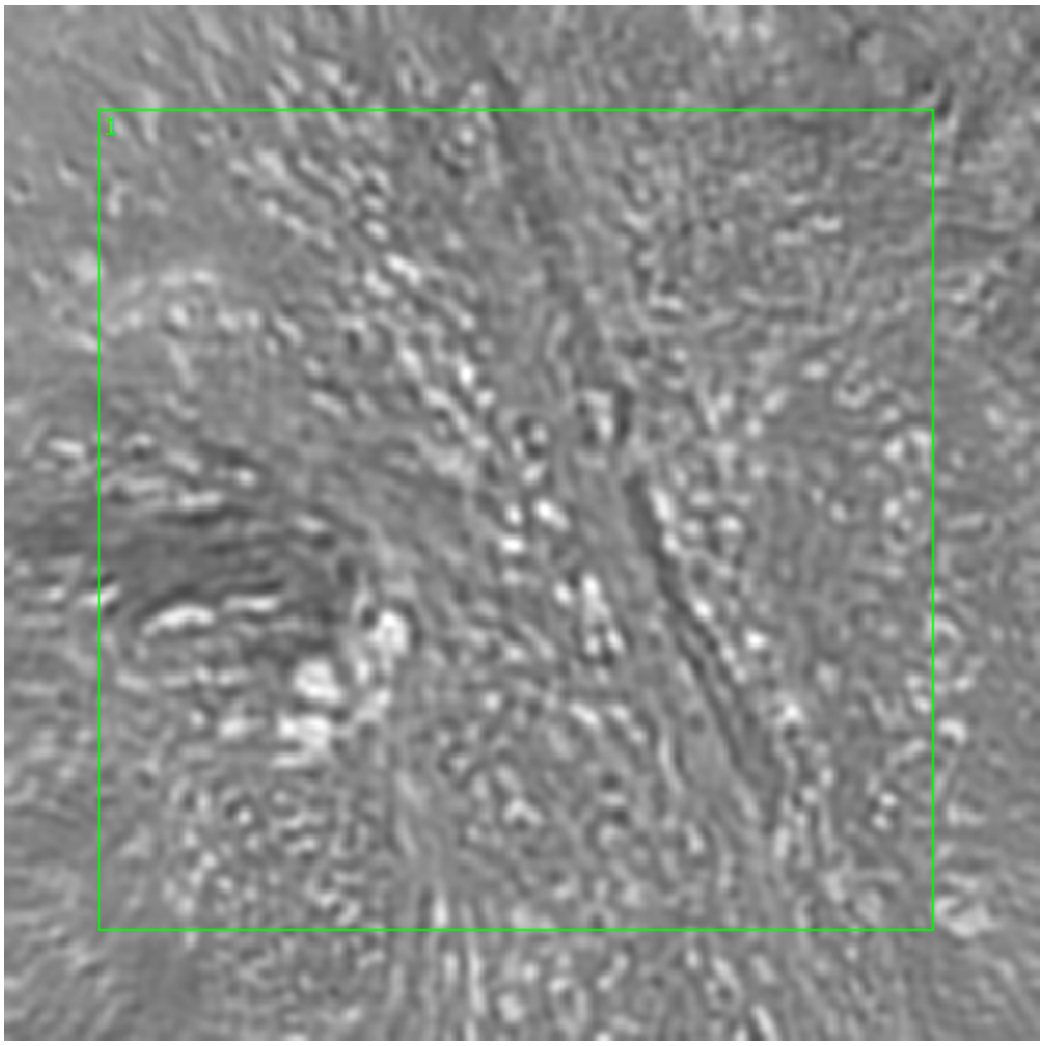



**Figure 2.46. Specimen 93-02-A251, Point F Regions of Interest**
**Region 1: lamina propria, squamous cell carcinoma**

The spectrum from the defined region of interest on Point F is depicted in Figure 2.47.

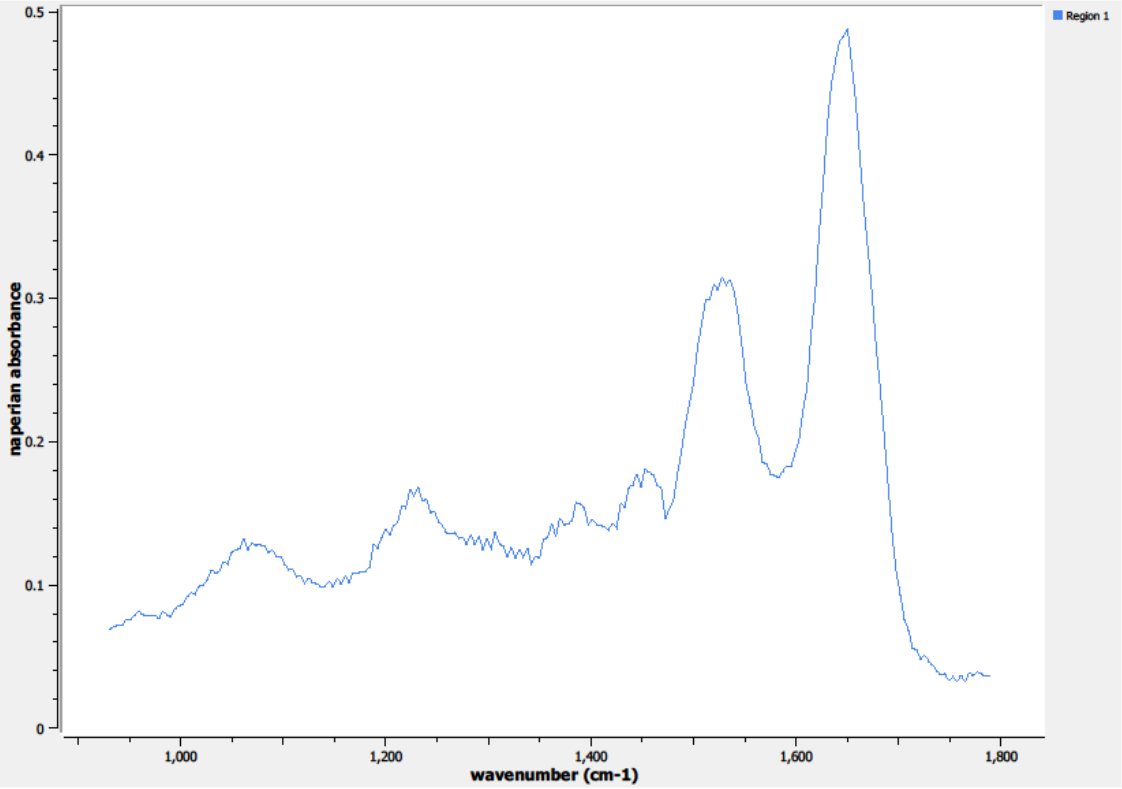



**Figure 2.47. Specimen 93-02-A251, Point F Spectra**

### 2.3.2.2. Specimen 91-10-A177

Specimen 91-10-A177 is a malignant cervical sample from a 29-year-old female. CHTN did not include a weight of the sample, a cervical location from which the specimen was removed, or the race of the patient from whom the specimen was removed. The quality control diagnostic provided by CHTN personnel identifies the specimen as papillary adenosquamous carcinoma of the cervix, indicating that the specimen was a protrusion on a glandular structure of the cervix. The tumor percentage was not provided with Specimen 91-10-A177.

An image of a sample from Specimen 91-10-A177 stained with H&E on a glass slide is shown in Figure 2.48. The dark purple throughout the image is indicative of high cell concentration typically associated with a tumor growth. The invasiveness of the tumor is evident throughout the entire tissue sample. The cervical glandular structures are also visible on the specimen.

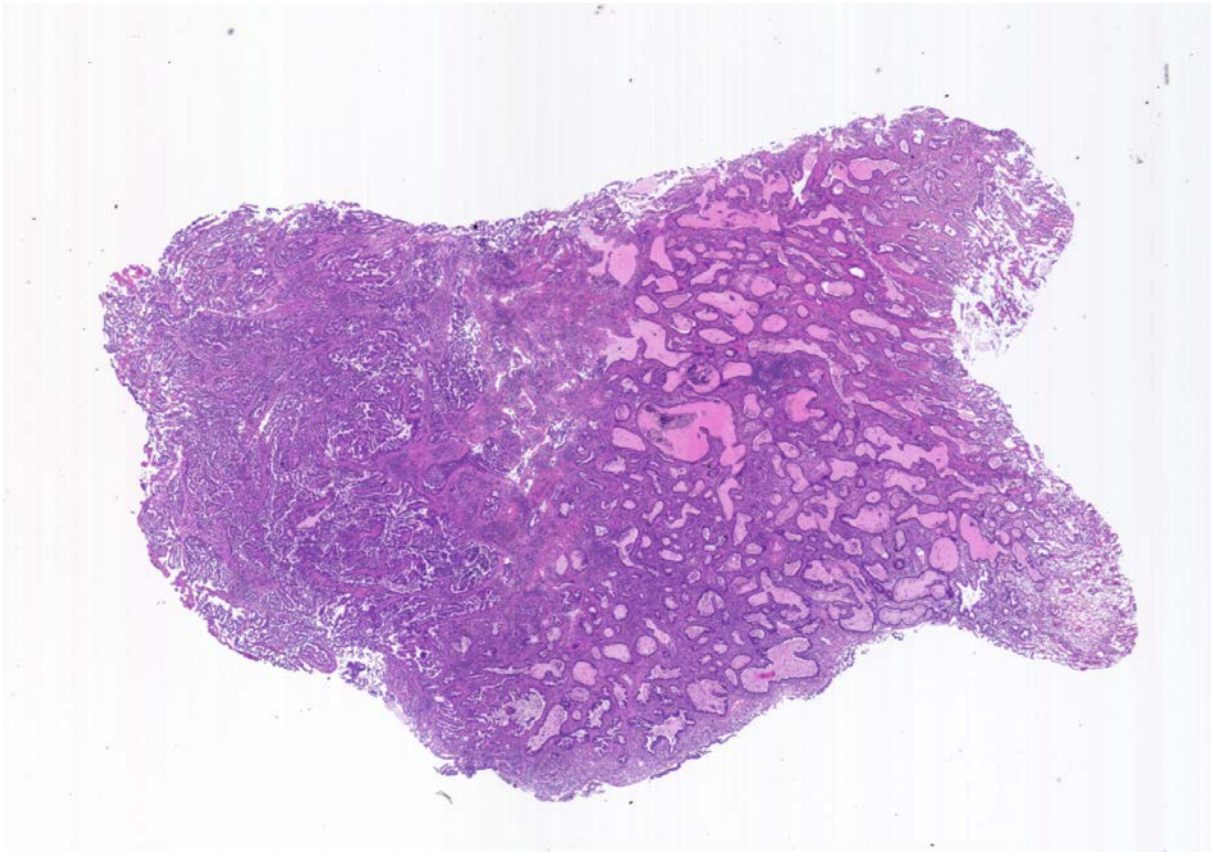


**Figure 2.48. Specimen 91-10-A177, H&E stain, papillary adenosquamous carcinoma of the cervix**

Figure 2.49 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags V and W depict the

spectral data collection points on Sample 91-10-A177, focusing on areas surrounding glandular structures.

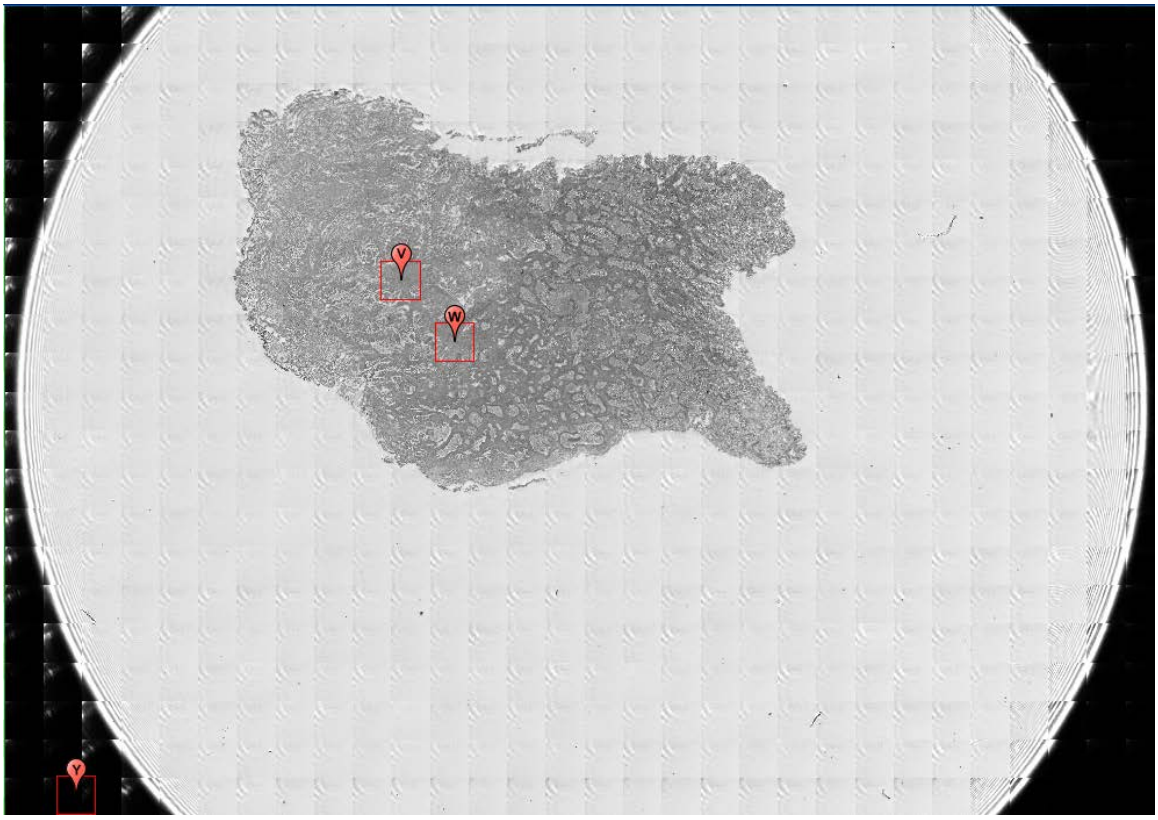



**Figure 2.49. Specimen 91-10-A177, cervical cancer tissue**

#### 2.3.2.2.1. Sample 91-10-A177, Point V

A 650 x 650 µm field-of-view image taken of the specimen at Point V (see Figure 2.50) was obtained from Spero's high-magnification objective. Point V is located on the interior of Sample 91-10-A177, in the lamina propria layer.

Regions of interest identified in Figure 2.50 are where the spectral information was actually collected on the sample. Regions 1 and 2 are both located

near glandular structures in the cervix, most likely mucous-secreting glands. Glandular structures are areas of interest for cervical carcinoma detection and identification.

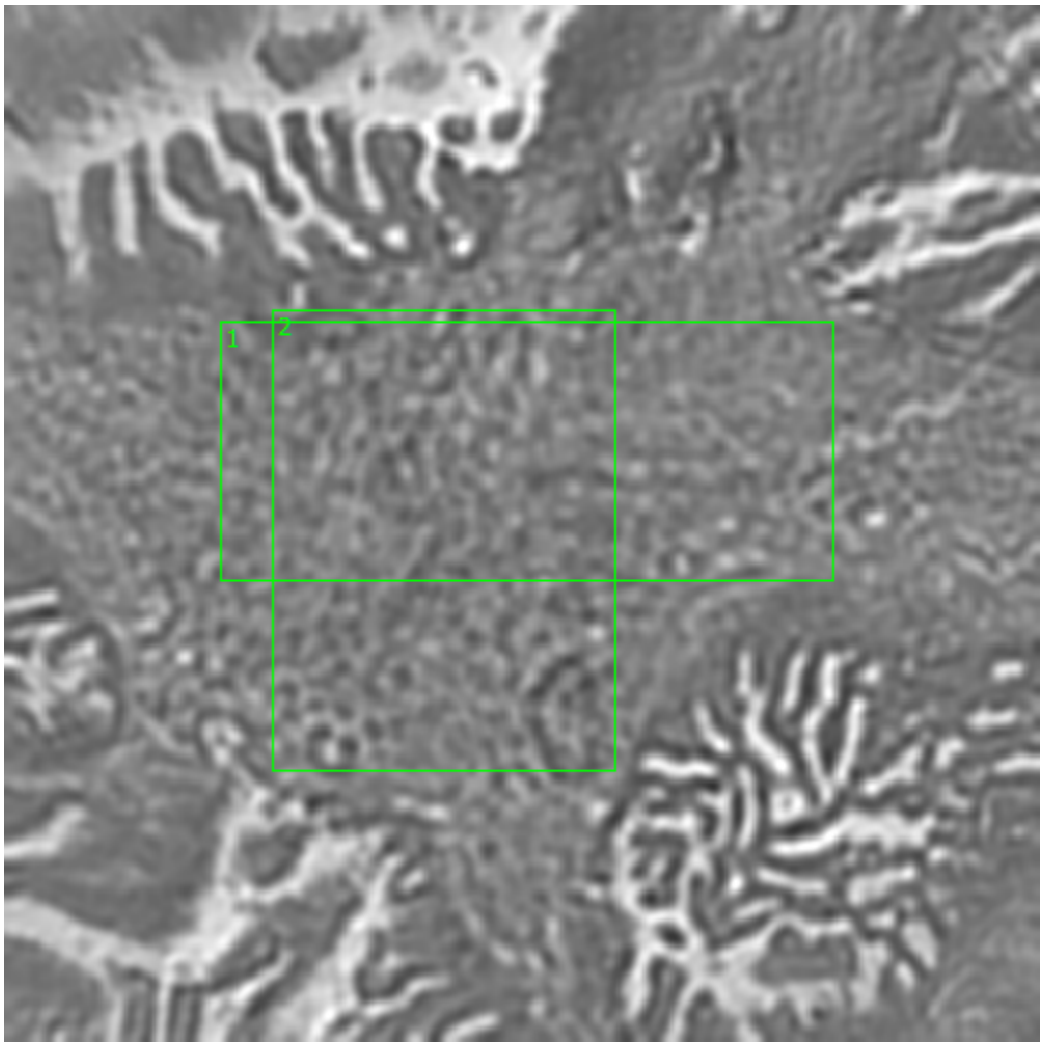



**Figure 2.50. Specimen 91-10-A177, Point V Regions of Interest**
**Regions 1, 2: mucous-secreting glands, papillary adenosqamous carcinoma**

The spectra from the two defined regions of interest on Point V are depicted in Figure 2.51.

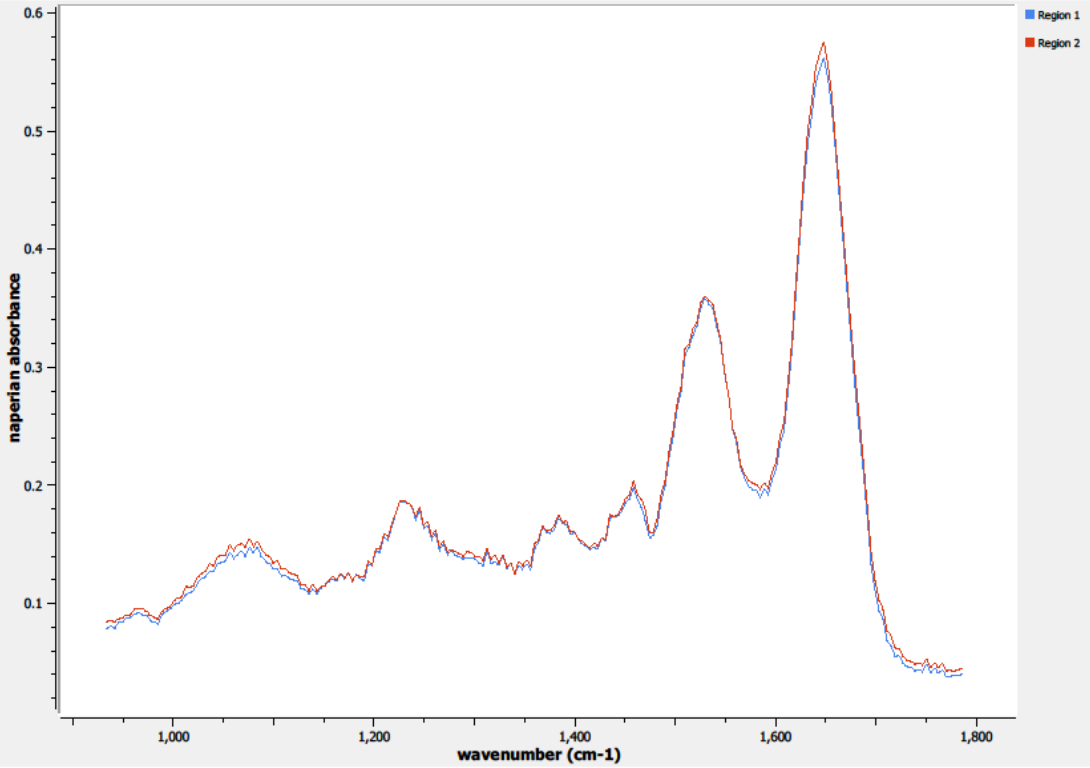



**Figure 2.51. Specimen 91-10-A177, Point V Spectra**

#### 2.3.2.2.2. Sample 91-10-A177, Point W

The high-magnification field-of-view image taken at Point W on Sample 91-10-A177 can be seen in Figure 2.52. Point W is located on the interior of the sample, in the lamina propria layer.

The spectral data collected on Point W were obtained from the regions of interest identified in Figure 2.52. Regions 1 and 2 are also both located near various glandular structures in the cervix.

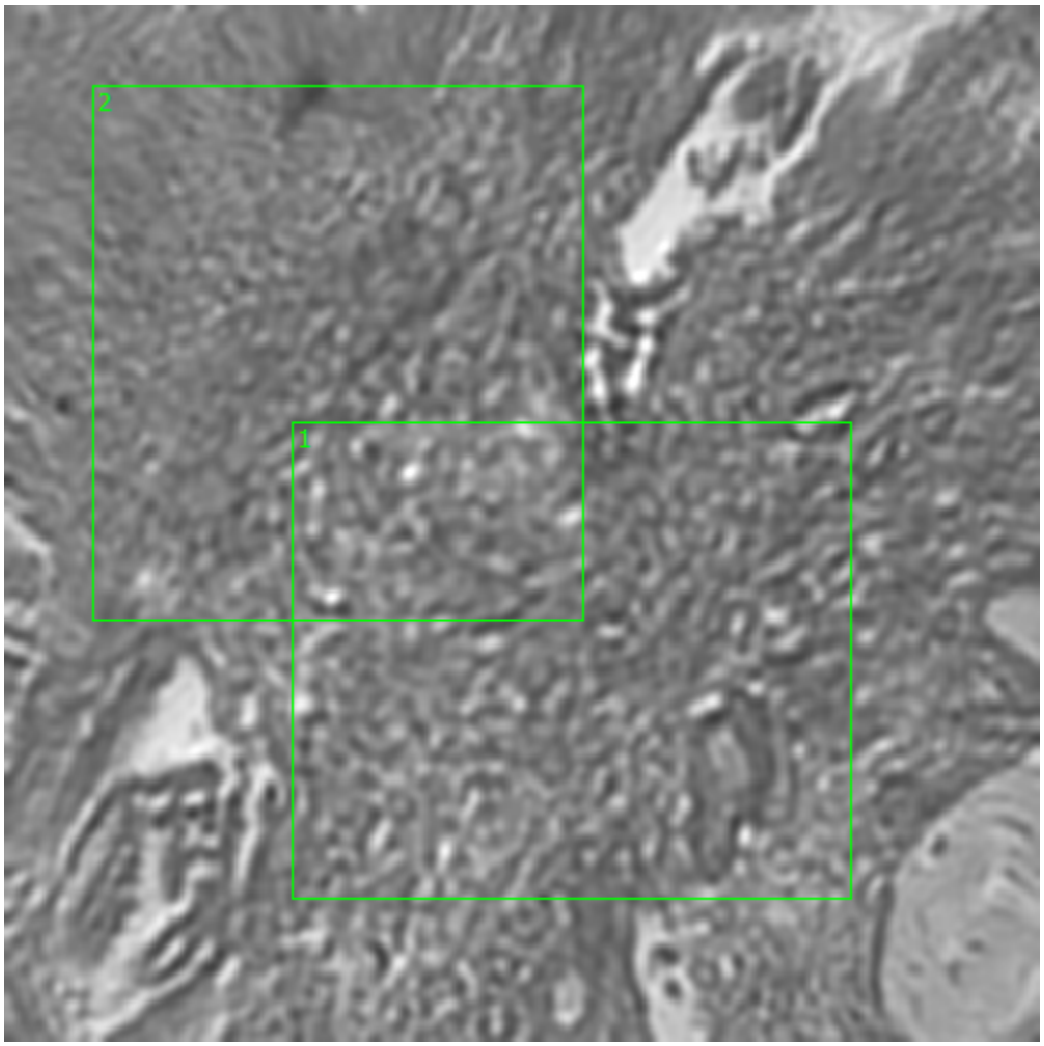



**Figure 2.52. Specimen 91-10-A177, Point W Regions of Interest**
**Regions 1, 2: mucous-secreting glands, papillary adenosquamous carcinoma**

The spectra from the two defined regions of interest on Point W are depicted in Figure 2.53.

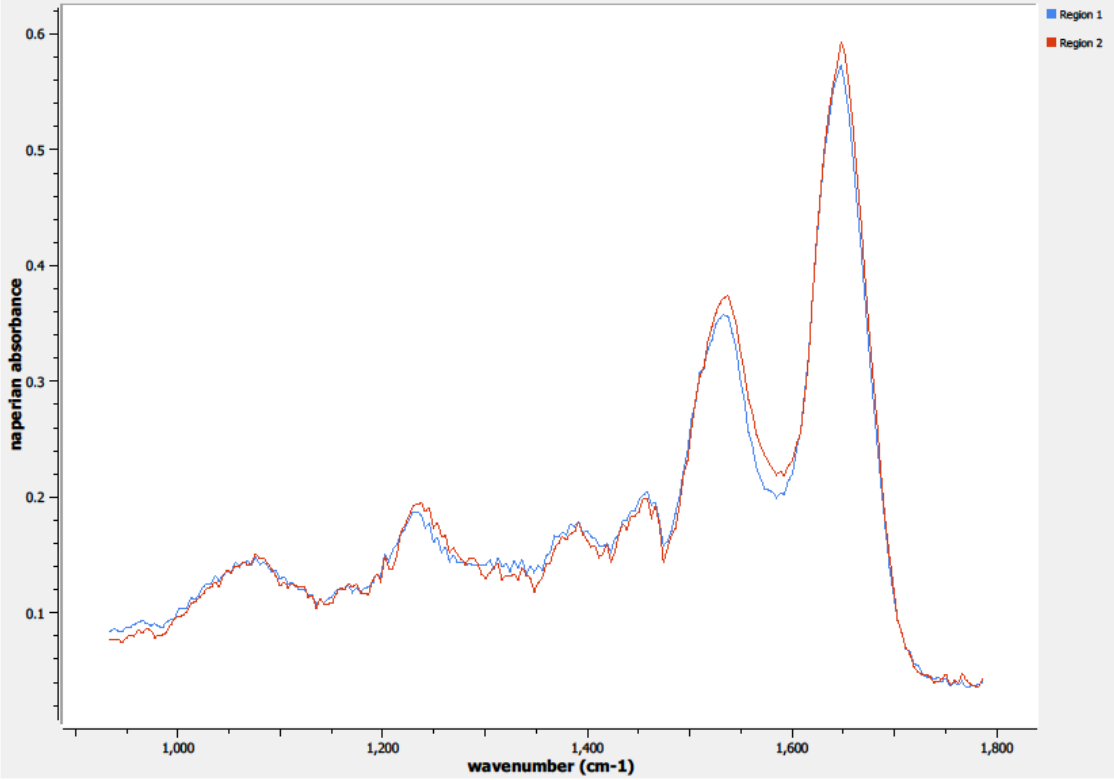



**Figure 2.53. Specimen 91-10-A177, Point W Spectra**

### 2.3.2.3. Specimen 10-09-A121a

Specimen 10-09-A121a is a malignant cervical sample from a 30-year-old white female. The specimen is a 0.65 gram primary malignant sample. CHTN did not include a cervical location from which the specimen was removed. The quality control diagnostic provided by CHTN personnel classifies the specimen as glassy cell

carcinoma of the cervix. The tumor percentage for Specimen 10-09-A121a is identified as 90% tumor.

Glassy cell carcinoma of the cervix is a rare and very aggressive form of cervical cancer that typically manifests in patients 10 years younger than other more common forms of cervical cancer. The cytoplasm conveys a glass-like appearance, which is where the name stems. Glassy cell carcinoma is a variant of adenosquamous carcinoma but is poorly differentiated, meaning that the tumor is fast-growing and aggressive.[56,57]

A 2002 NIH study states that although glassy cell carcinoma only accounts for approximately 5% of cervical cancer, the survival rate is only around 55% due to the aggressive nature of the tumor type. Some of the prognostic histological characteristics in squamous cell carcinoma can be utilized for glassy cell carcinoma predictions; however, the predictive qualities of these squamous cell features are not nearly as effective on glassy cell carcinoma.[56,57]

An image of a sample from Specimen 10-09-A121a stained with H&E on a glass slide is shown in Figure 2.54. The dark purple throughout the image is indicative of high cell concentration typically associated with a tumor growth. The invasiveness of the tumor is evident throughout the entire tissue sample.

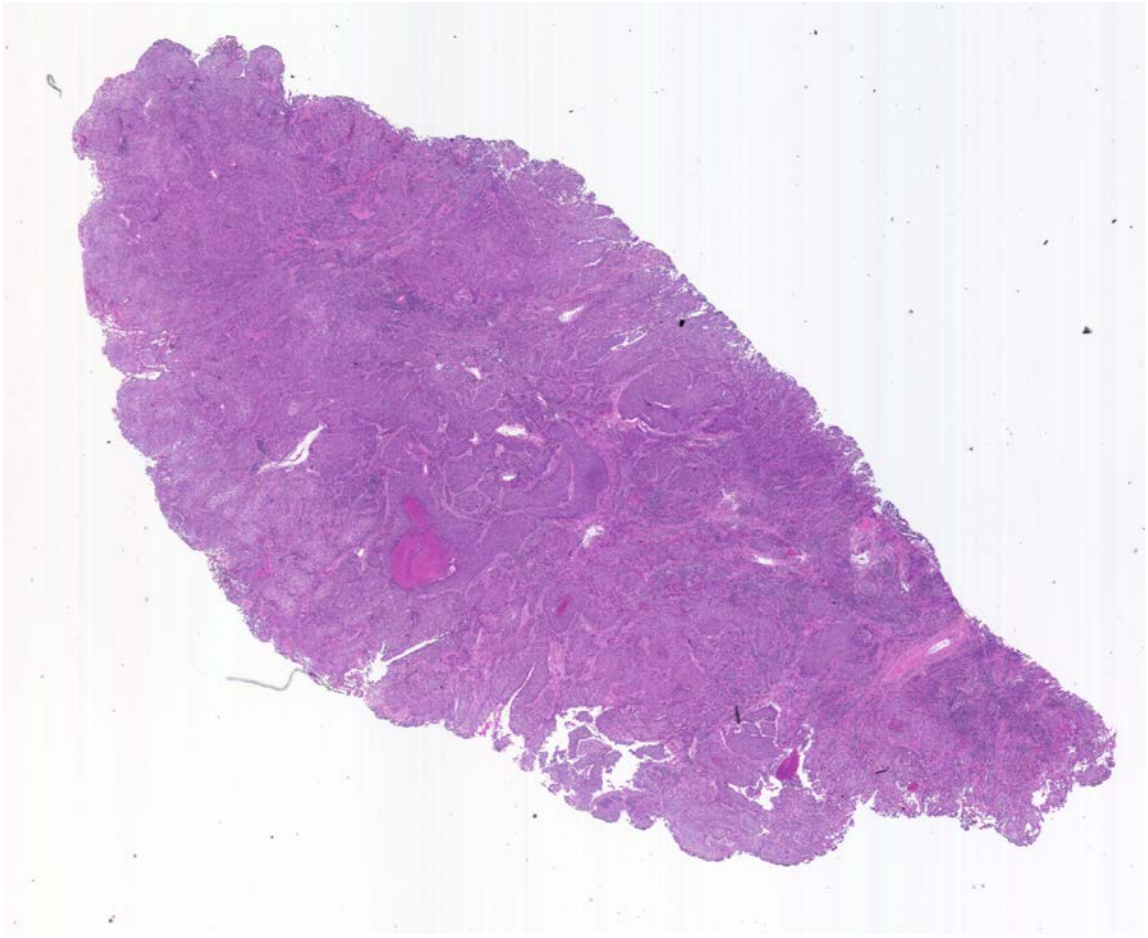


**Figure 2.54. Specimen 10-09-A121a, H&E stain, glassy cell carcinoma of the cervix**

Figure 2.55 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags J, K, and L depict three different spectral data collection points in the lamina propria layer on Sample 10-09-A121a.

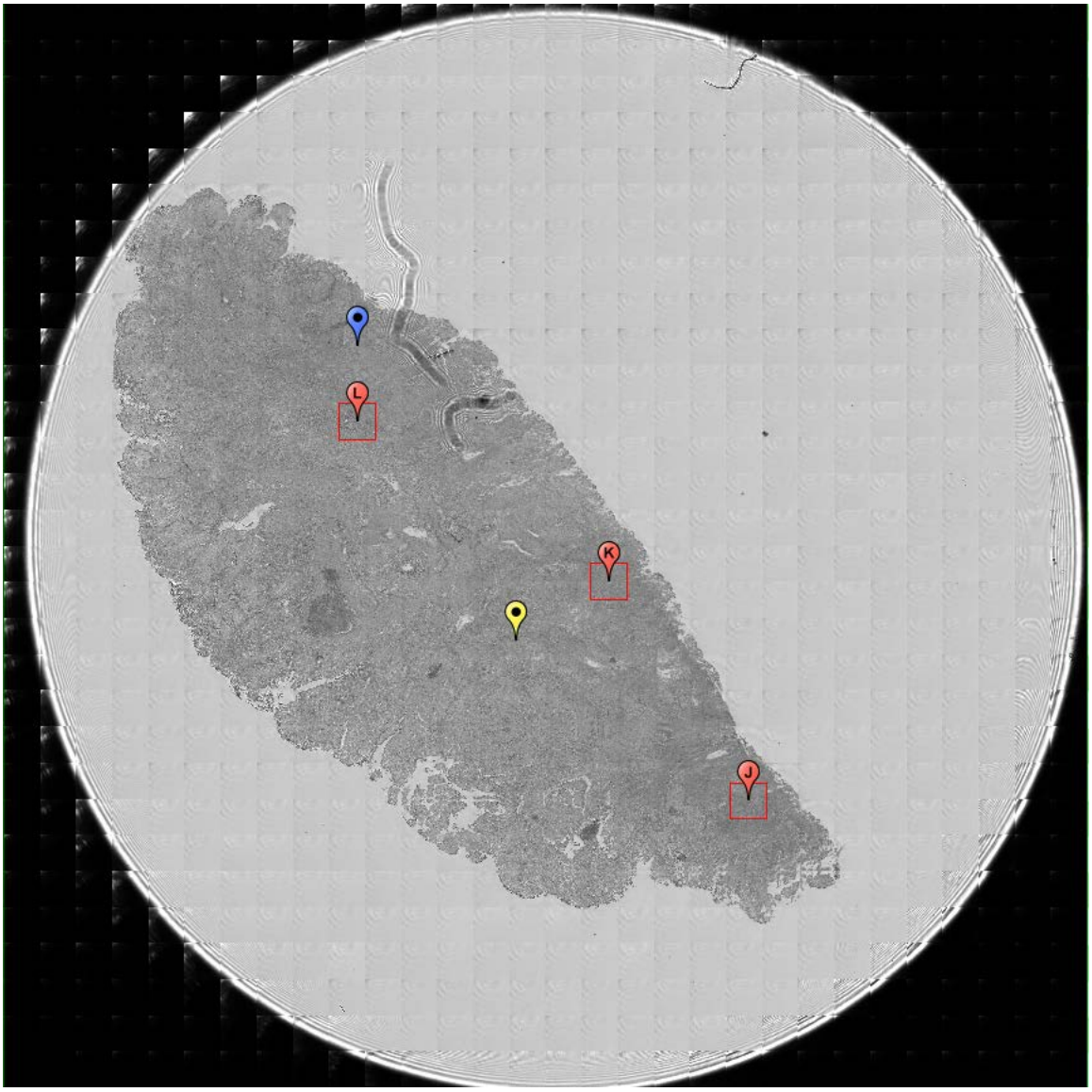



**Figure 2.55. Specimen 10-09-A121a, cervical cancer tissue**

#### 2.3.2.3.1. Sample 10-09-A121a, Point J

Figure 2.56 shows a 650 x 650 µm field-of-view image taken of Sample 10-09-A121a at Point J. This image was obtained by utilizing Spero's high-magnification objective. Point J is located on the interior of Sample 10-09-A121a, in

the lamina propria layer. Only one region of interest was needed to capture the spectral information at Point J.

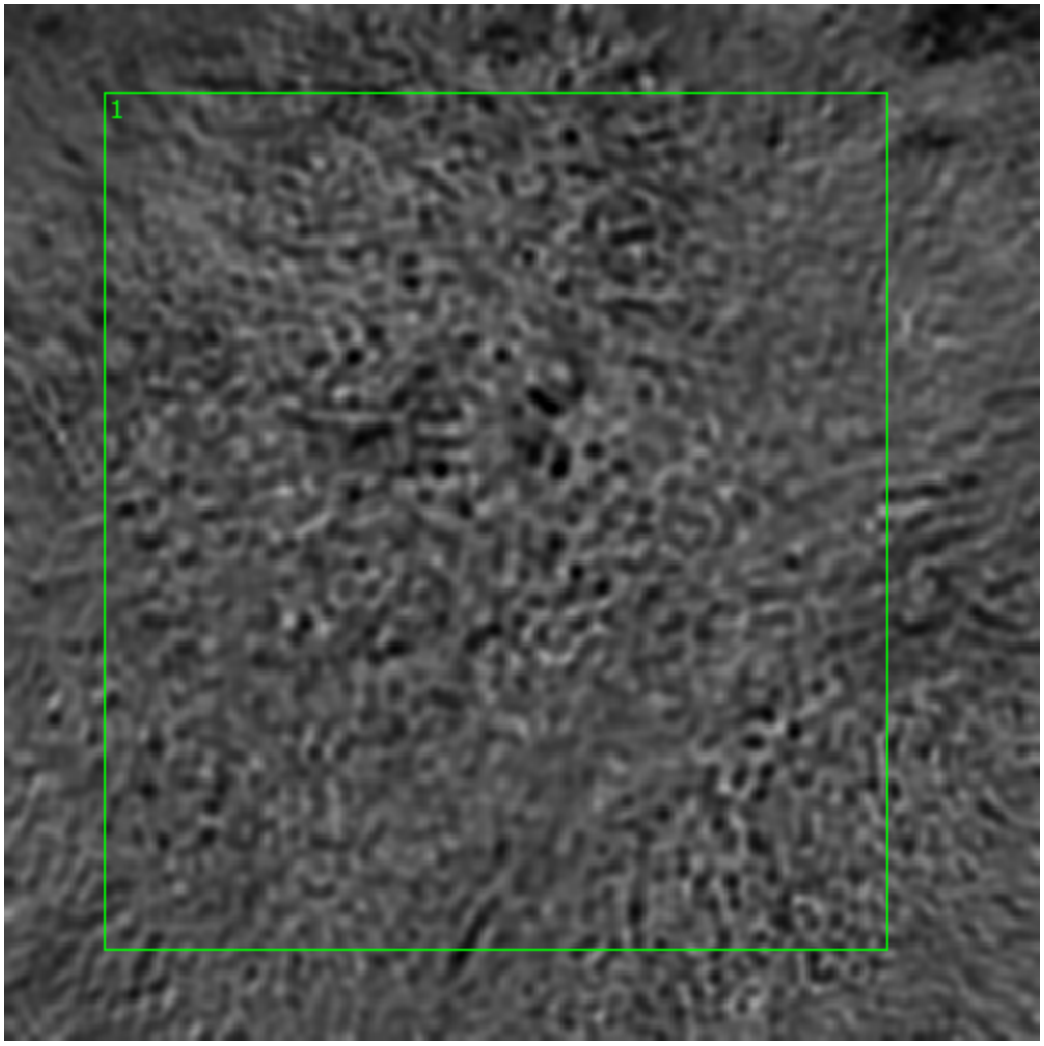



**Figure 2.56. Specimen 10-09-A121a, Point J Regions of Interest**
**Region 1: lamina propria, glassy cell carcinoma**

The spectrum from the defined region of interest on Point J is depicted in Figure 2.57.

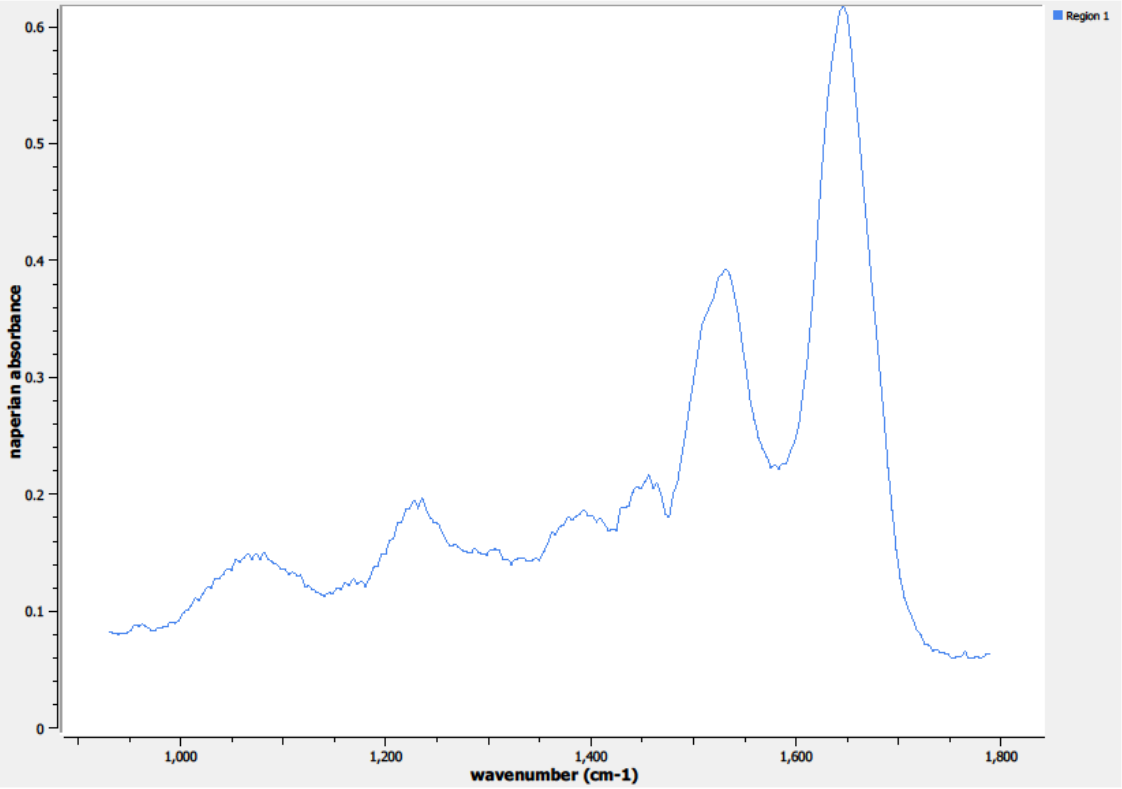



**Figure 2.57. Specimen 10-09-A121a, Point J Spectra**

#### 2.3.2.3.2. Sample 10-09-A121a, Point K

A 650 x 650 µm field-of-view image taken of Sample 10-09-A121a at Point K is seen in Figure 2.58. This image was obtained via Spero's high-magnification objective. Point K is located on the interior of Sample 10-09-A121a, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at Point K.

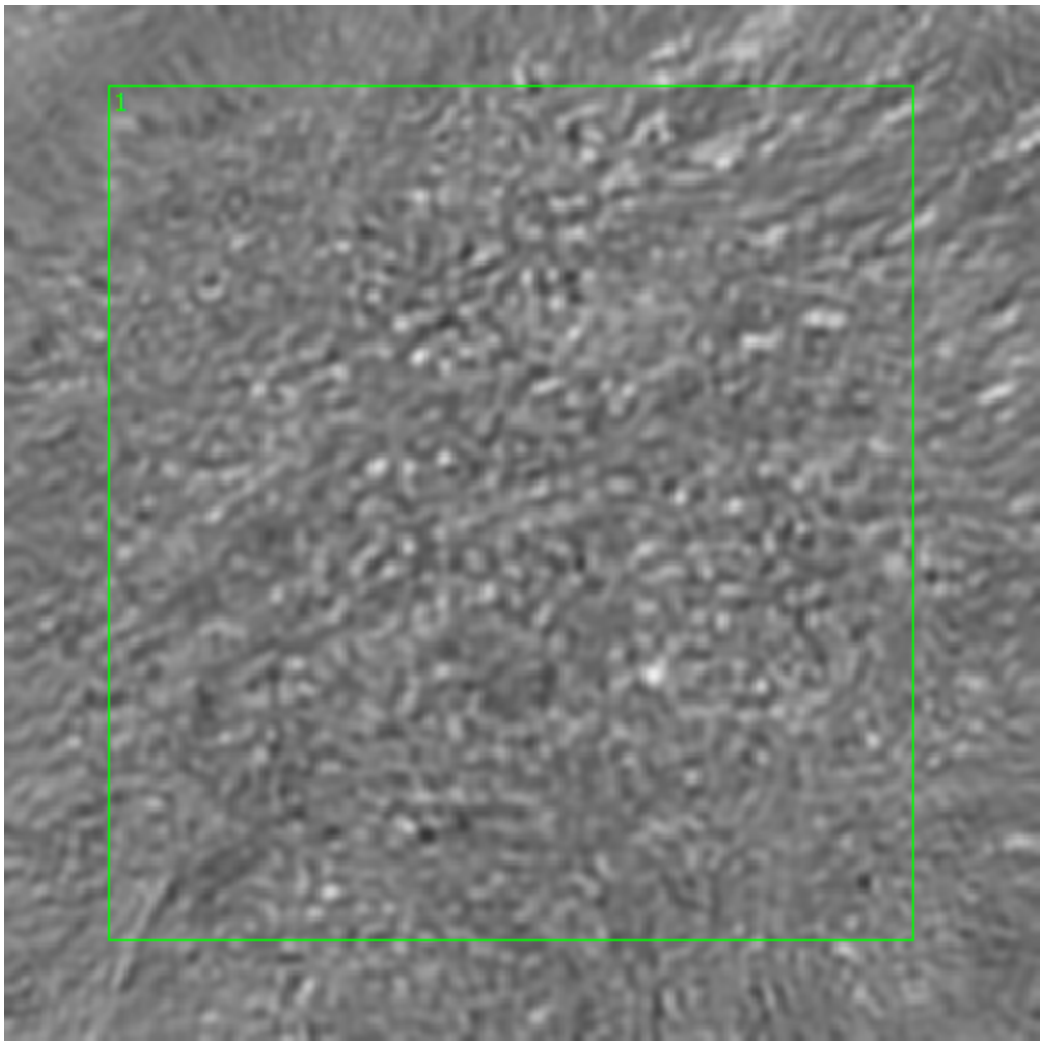



**Figure 2.58. Specimen 10-09-A121a, Point K Regions of Interest**
**Region 1: lamina propria, glassy cell carcinoma**

The spectrum from the defined region of interest on Point K is depicted in Figure 2.59.

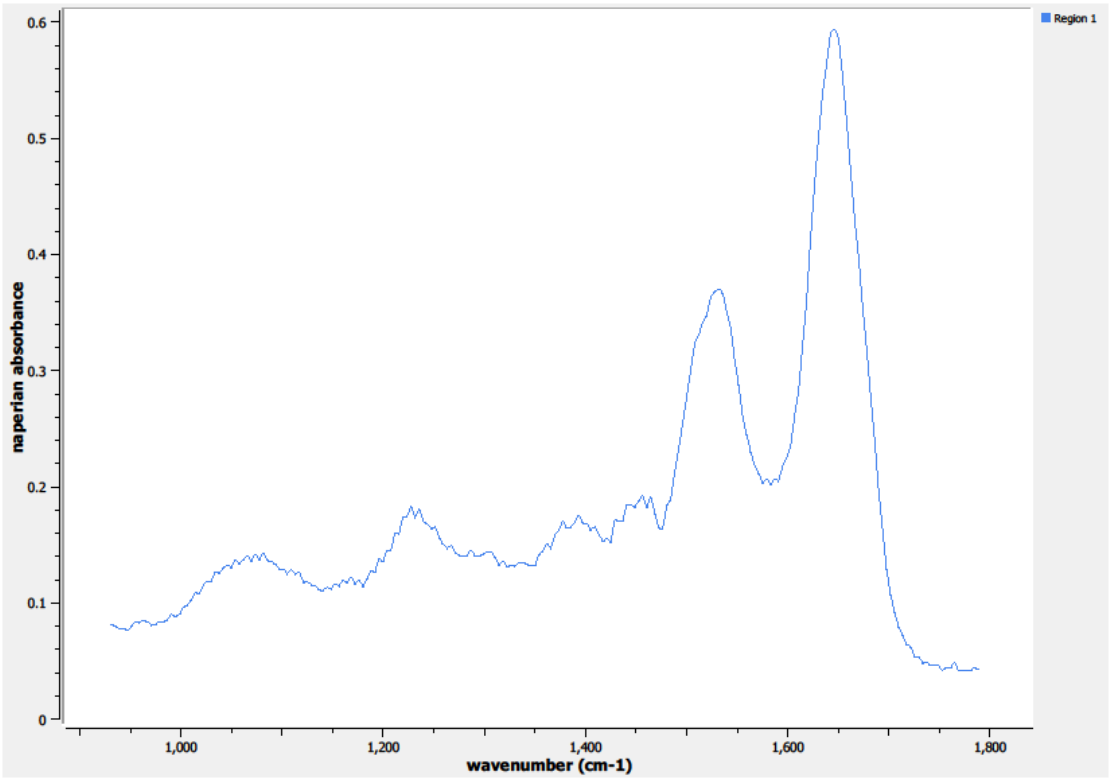



**Figure 2.59. Specimen 10-09-A121a, Point K Spectra**

#### 2.3.2.3.3. Sample 10-09-A121a, Point L

A 650 x 650 µm field-of-view image taken of Sample 10-09-A121a at Point L is seen in Figure 2.60. This image was obtained by using Spero's high-magnification objective. Point L is located on the interior of Sample 10-09-A121a, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at Point L.

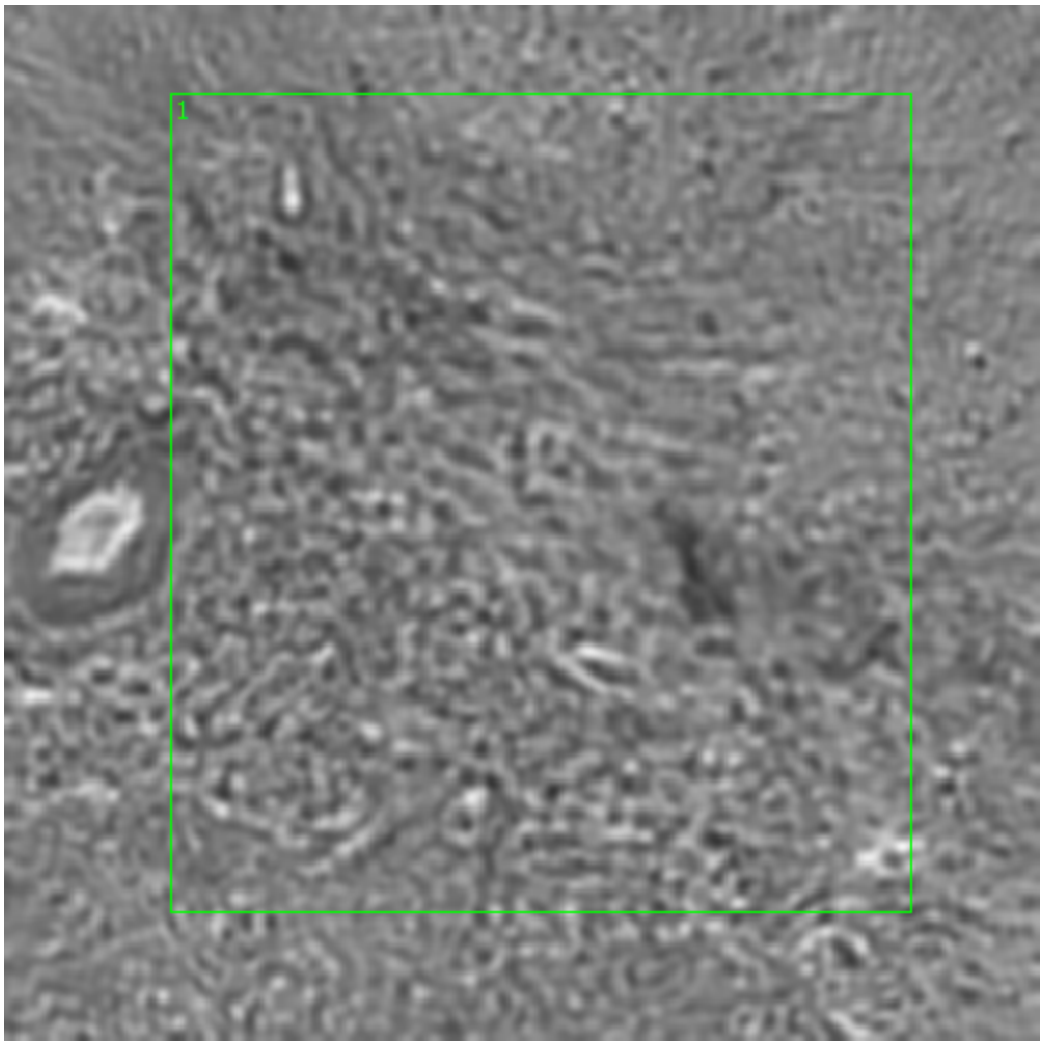



**Figure 2.60. Point L Regions of Interest**
**Region 1: lamina propria, glassy cell carcinoma**

The spectrum from the defined region of interest on Point L is depicted in Figure 2.61.

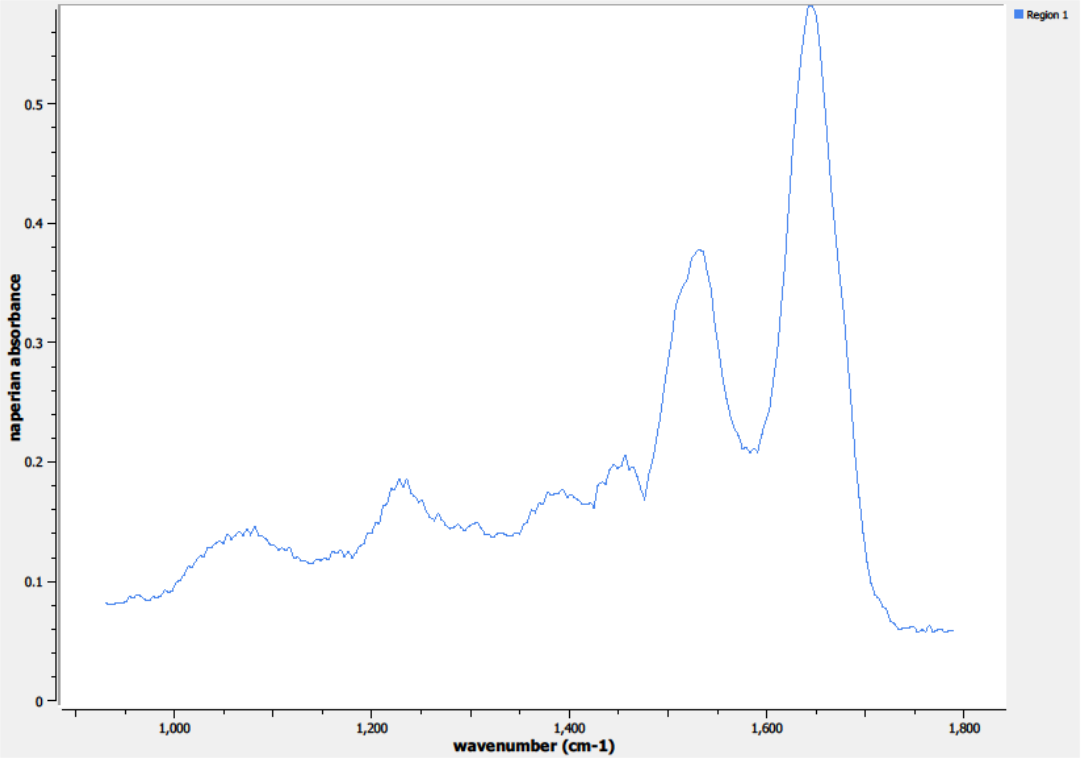



**Figure 2.61. Point L Spectra**

### 2.3.2.4. Specimen 06-02-A042b

Specimen 06-02-A042b is a malignant cervical sample from a 43-year-old white female. The specimen is a 0.61 gram primary malignant sample. CHTN did not include a cervical location from which the specimen was removed. This specimen is another example of glassy cell carcinoma of the cervix. The quality control diagnostic provided by CHTN personnel identifies the specimen as a glassy cell primary malignant adenosquamous carcinoma of the cervix. The tumor percentage for Specimen 06-02-A042b is identified as 50% tumor.

An image of a sample from Specimen 06-02-A042b stained with H&E on a glass slide is shown in Figure 2.62. The dark purple throughout the image is indicative of high cell concentration typically associated with a tumor growth. The invasiveness of the poorly differentiated glassy cell tumor growth is evident throughout the entire tissue sample.

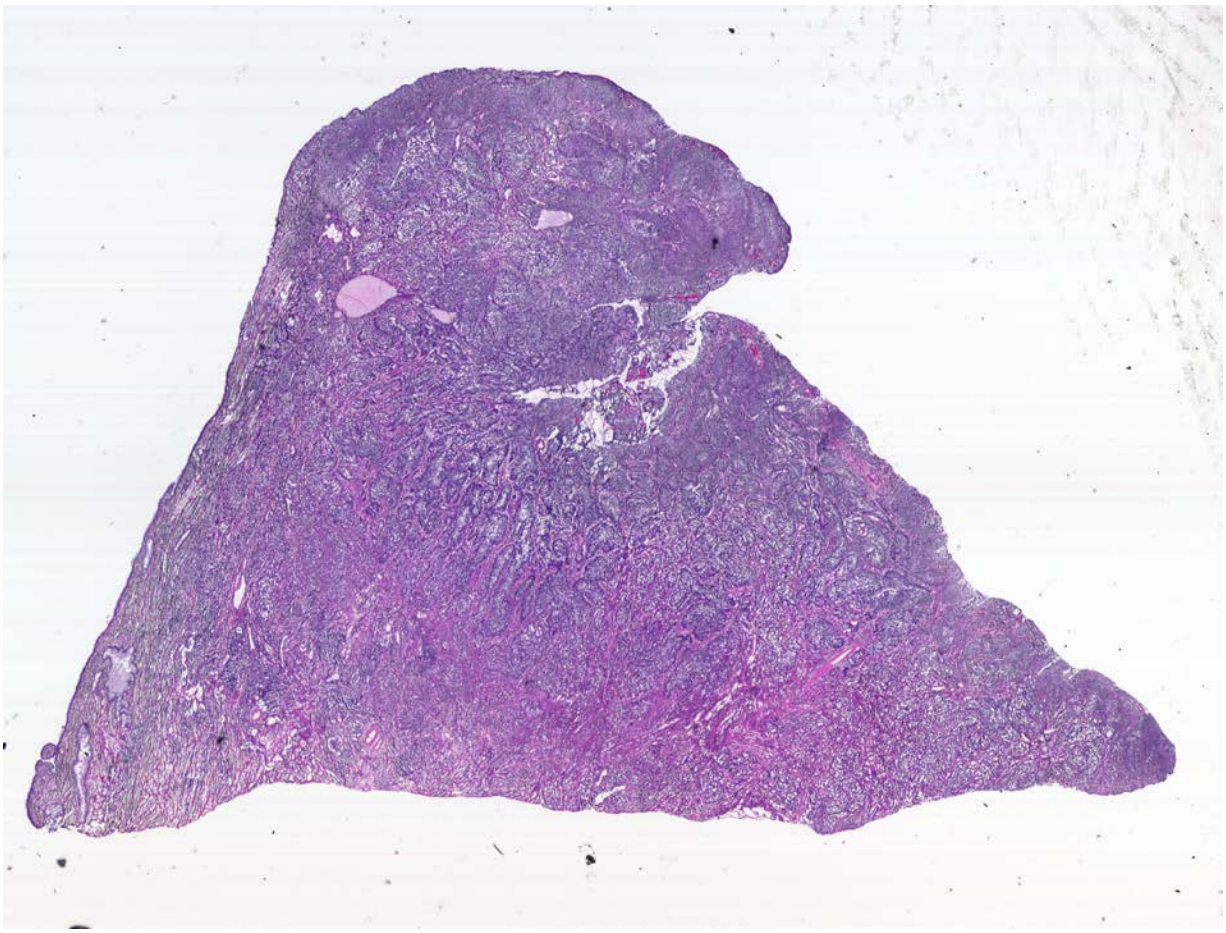


**Figure 2.62. Specimen 06-02-A042b, H&E stain, adenosquamous carcinoma of the cervix, glassy cell**

Figure 2.63 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags P and W depict two

different spectral data collection points in the lamina propria layer on Sample 06-02-A042b. Both flag locations targeted darker areas of high cell concentration on the sample.

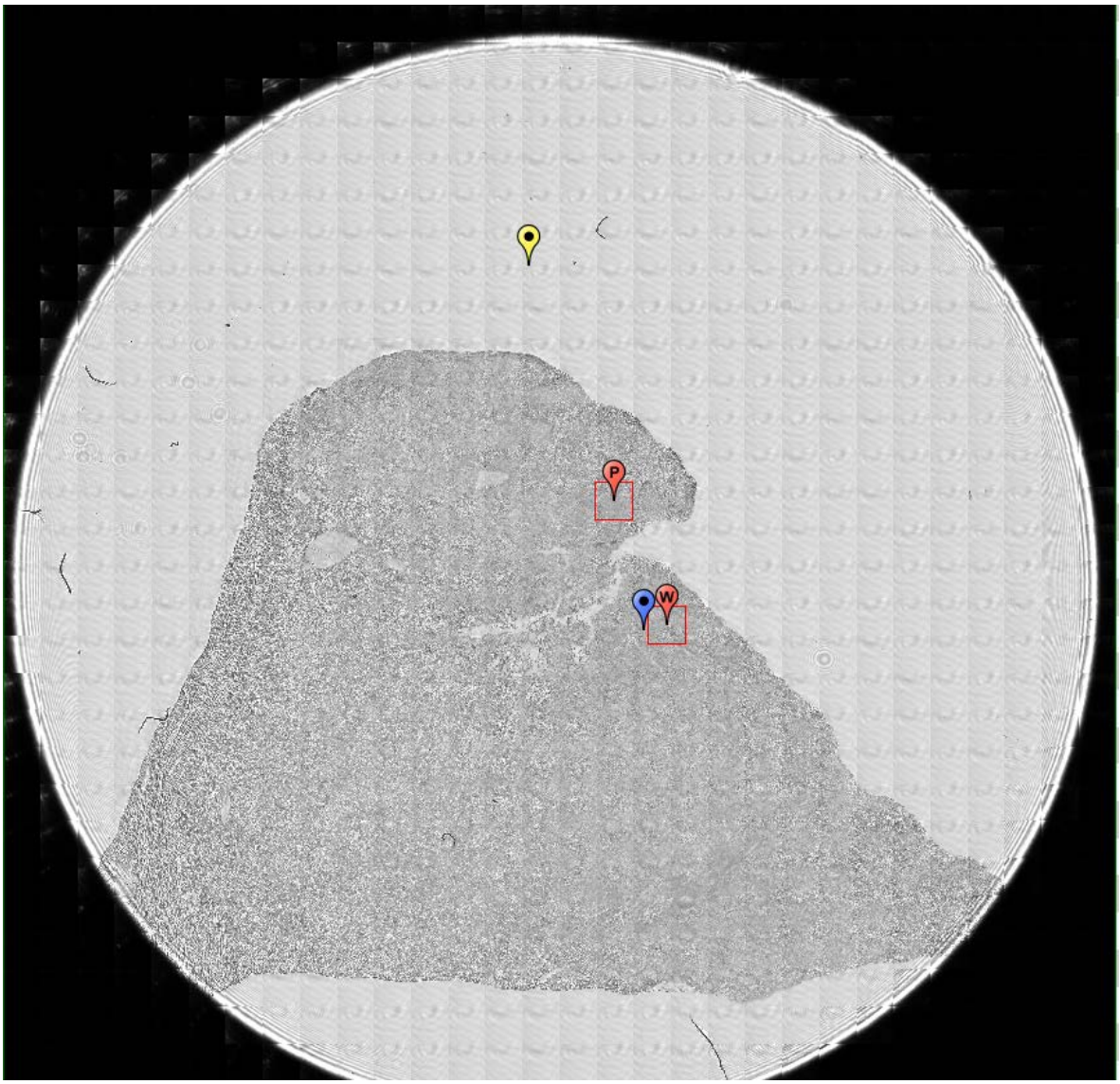



**Figure 2.63. Specimen 06-02-A042b, cervical cancer tissue**

#### 2.3.2.4.1. Sample 06-02-A042b, Point P

Figure 2.64 shows a 650 x 650 μm field-of-view image taken of Sample 06-02-A042b at Point P. This image was obtained by utilizing Spero's high-magnification objective. Point P is located on the interior of Sample 06-02-A042b, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at Point P.

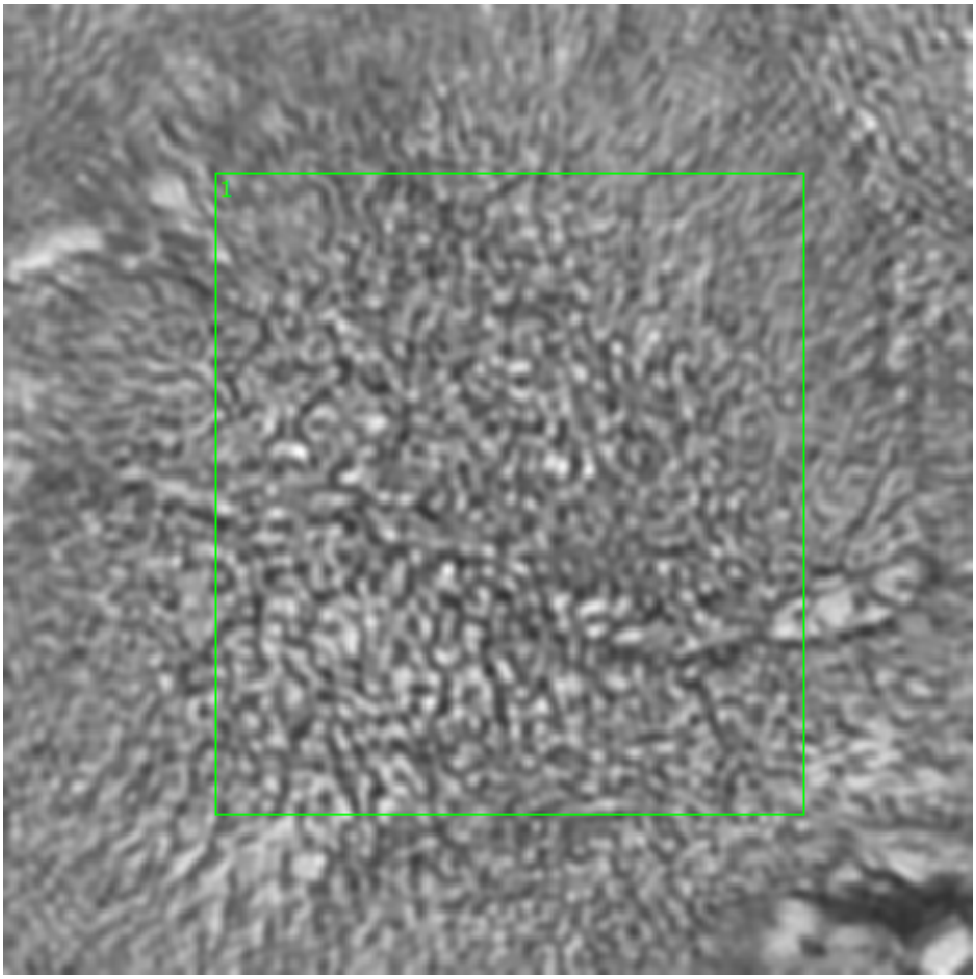


**Figure 2.64. Sample 06-02-A042b, Point P Regions of Interest**
**Region 1: lamina propria, adenosquamous glassy cell carcinoma**

The spectrum from the defined region of interest on Point P is depicted in Figure 2.65.

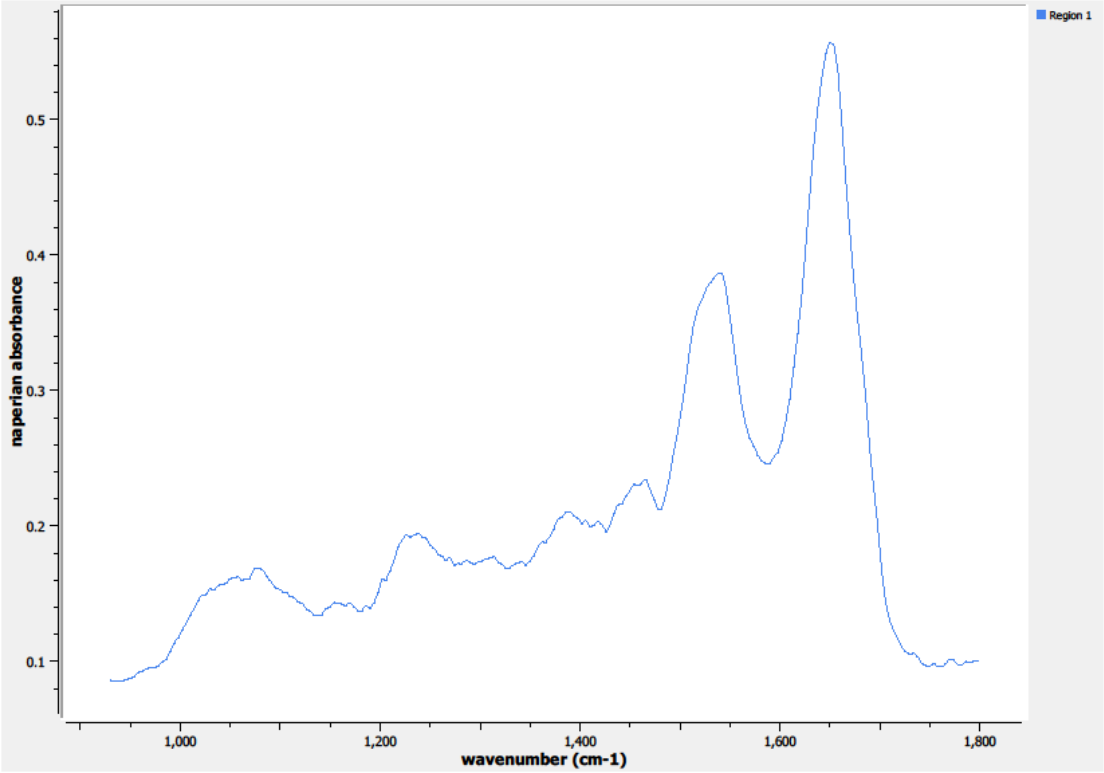



**Figure 2.65. Sample 06-02-A042b, Point P Spectra**

#### 2.3.2.4.2. Sample 06-02-A042b, Point W

Figure 2.66 shows an image taken of Sample 06-02-A042b at Point W by utilizing Spero's high-magnification objective. Point W is located on the interior of Sample 06-02-A042b, in the lamina propria layer. Only one region of interest was needed to capture the spectral information at Point W.

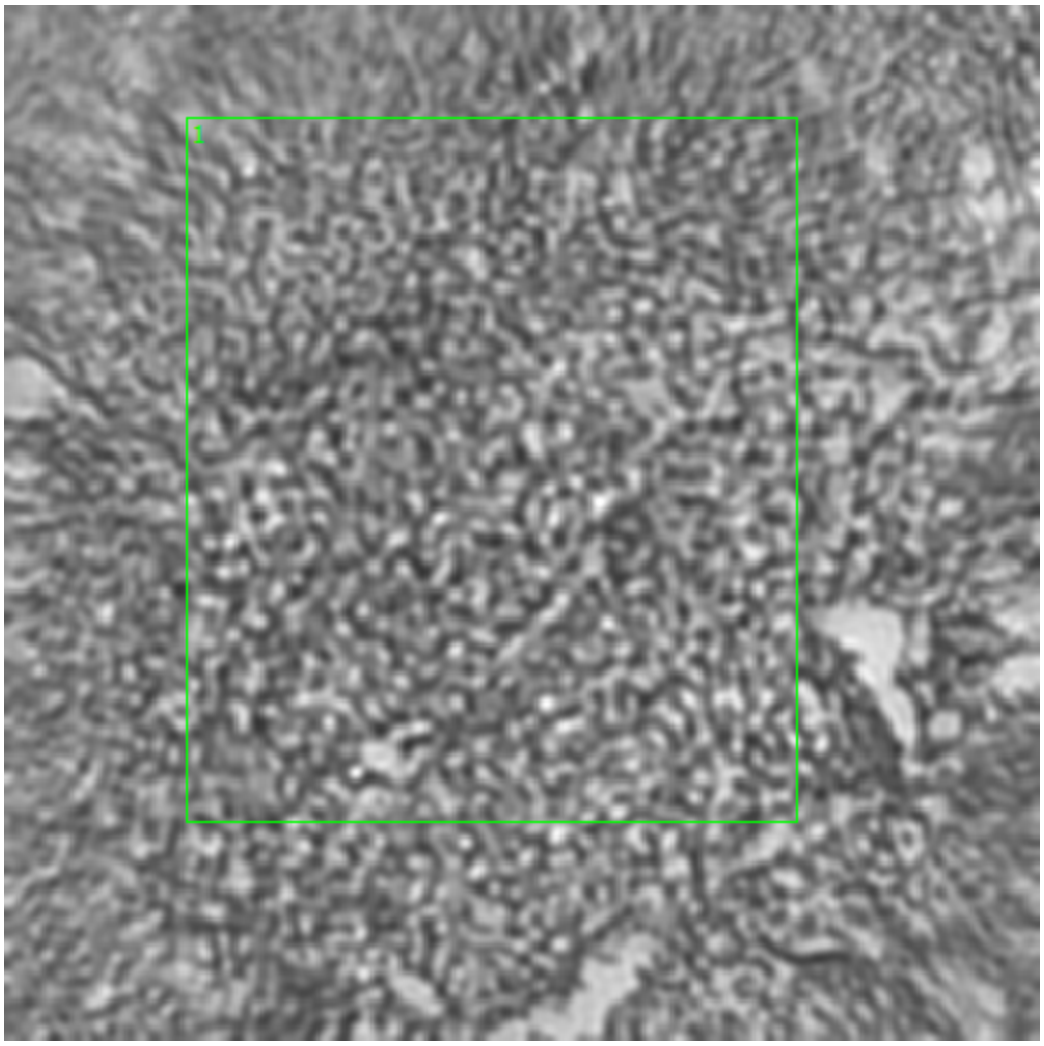



**Figure 2.66. Sample 06-02-A042b, Point W Regions of Interest**
**Region 1: lamina propria, adenosquamous glassy cell carcinoma**

The spectrum from the defined region of interest on Point W is depicted in Figure 2.67.

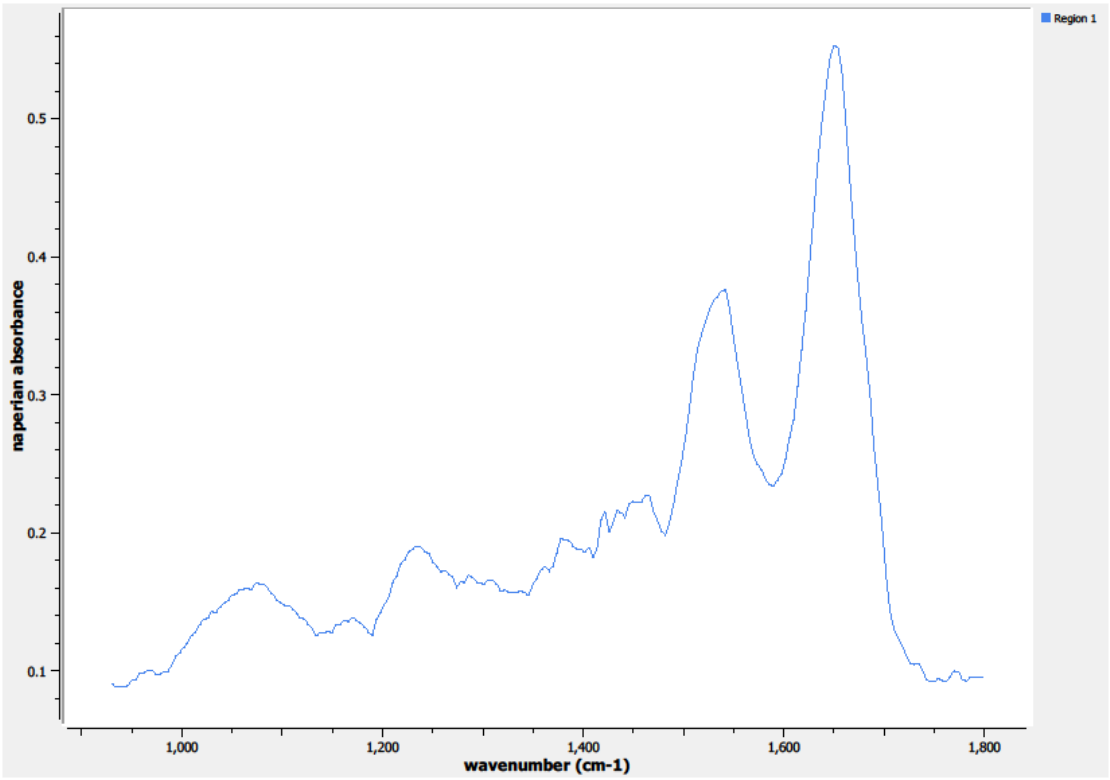



**Figure 2.67. Sample 06-02-A042b, Point W Spectra**

## 2.4. Analysis

We used the two-tailed T-test to evaluate statistical differences in our results since deviations could be in either direction. Statistical differences were considered to be significant at $p < 0.05$. All spectra and analysis figures shown were obtained from the mean spectra of the individual specimen groups (normal or malignant, endocervix or ectocervix, etc.) and each tissue layer studied within the groups

(epithelium, basal, lamina propria). Spectra and analysis values shown represent the entire experiment.

### 2.4.1. Intensity Ratio Parameter Analysis

Prominent absorbance peak values in the "fingerprint" region were used for intensity ratio parameter analysis. The peaks were chosen by a combination of observation and taking into account peaks evaluated from the literary FTIR review described in Chapter 1. All ratios utilized the amide I peak (1654 $cm^{-1}$) as the denominator and all other selected peak values as the numerator.

The observed peak that is the most dominant in absorbance is the protein amide I peak, which is mainly attributed to C=O stretching vibration. The protein amide II peak at 1544 $cm^{-1}$ is also significant, which arises from C-N stretching vibration coupled with N-H bending vibration.

Both bands at 1454 $cm^{-1}$ and 1402 $cm^{-1}$ arise from CH3 bending modes of methyl proteins and lipids; 1454 $cm^{-1}$ is attributed to asymmetrical bending and 1402 $cm^{-1}$ is attributed to symmetrical bending. The band at 1338 $cm^{-1}$ is much less prominent that others, but it does show increased absorption in some tissue groupings and no increased absorption in others. This band is in the peptide group and is evident of normal mode symmetry and C-H vibration.

Amide III is another common protein peak used for investigations, which is around 1240 $cm^{-1}$. This absorption band is mainly due to asymmetrical phosphate ($PO_2^-$) stretching modes of nucleic acids and vibrations of collagen. The absorption

peak at 1080 $cm^{-1}$ is also impacted by nucleic acids and vibration modes in collagen; in this band, though, the impact is due to symmetrical stretching of $PO_2^-$. The final absorption band we evaluated was a weak absorption sometimes seen at 1025 $cm^{-1}$. The absorption at this band is attributed mainly to the C-O vibrational mode and the C-O stretching/bending of glycogen. Again, similar to 1338 $cm^{-1}$, an absorption peak was not present in all our specimen tissue groupings.

#### 2.4.1.1. Normal Endocervix

The first comparison which we investigated was evaluating among epithelium, basal, and lamina propria layers on normal endocervical specimens to see if we could successfully discriminate. Variances are depicted in the infrared spectra seen in Figure 2.68.

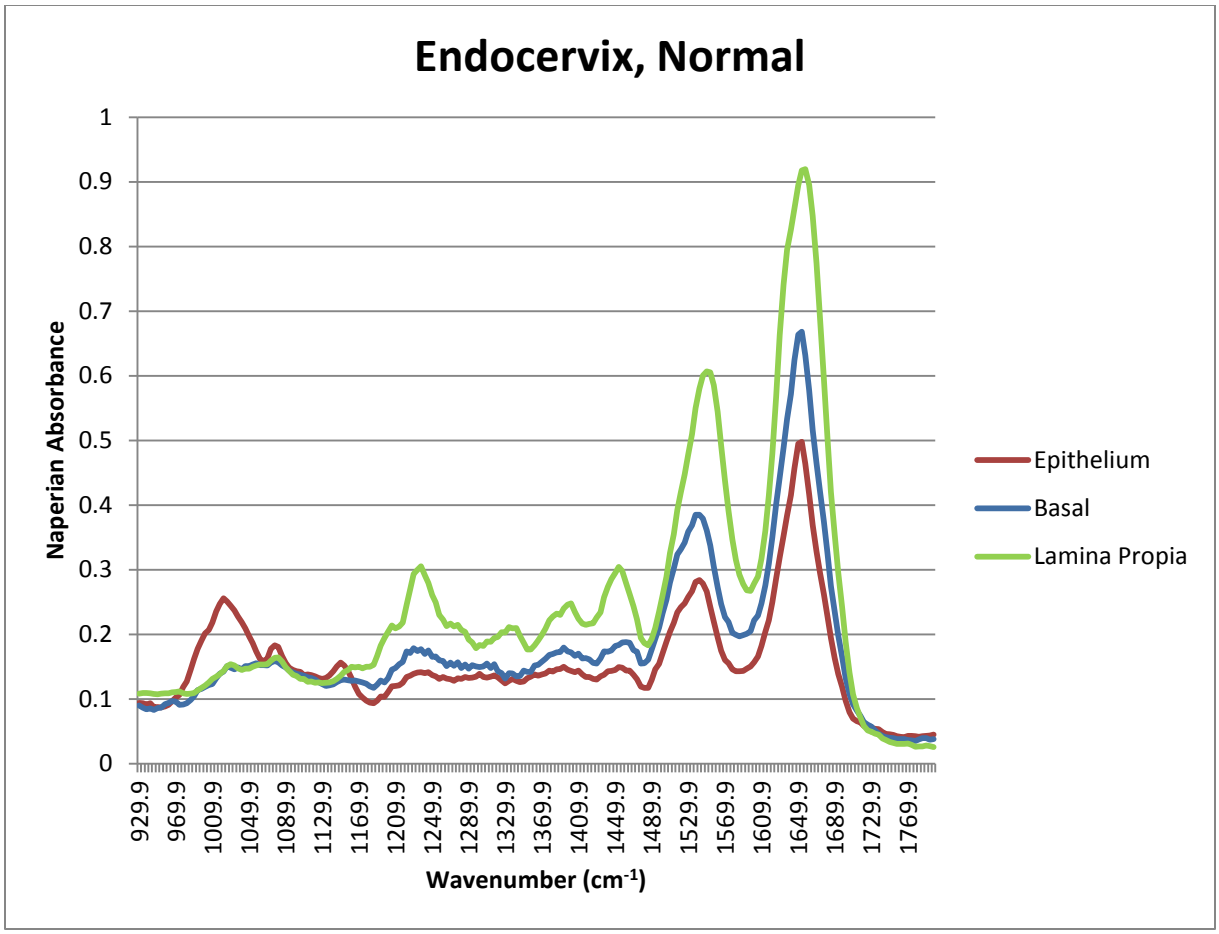



**Figure 2.68. IR spectra of average of normal endocervical specimens, comparing epithelium, basal, and lamina propria layers**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 2.68 are shown in Figure 2.69. The plot shows the differences at peak intensity values of various bands in the spectra specific to the three layers being assessed.

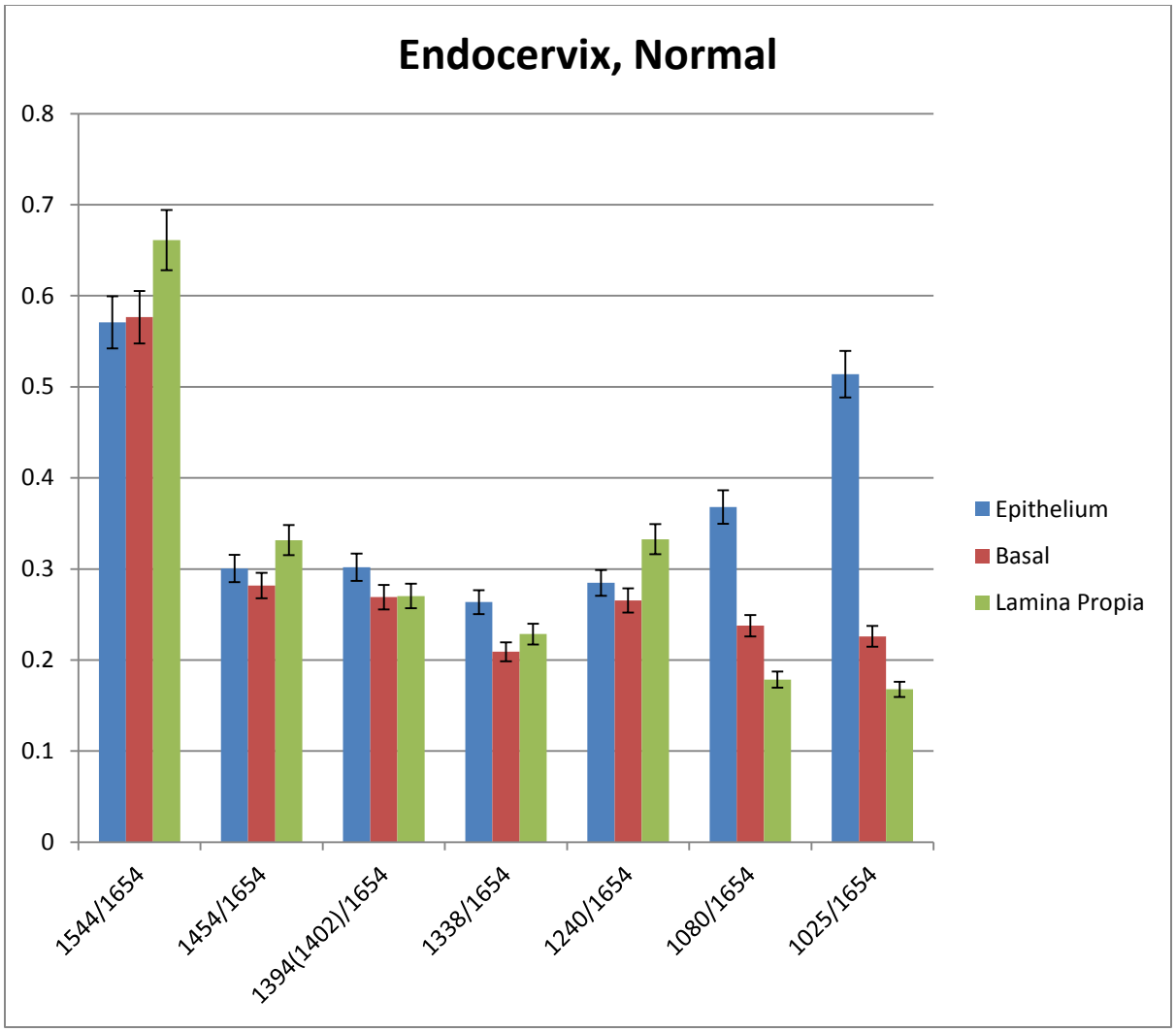



**Figure 2.69. Mean intensity ratio parameter of normal endocervical specimens at different bands throughout the fingerprint region for epithelium, basal, and lamina propria layers**

Five of the seven intensity ratio groupings show statistically significant results for normal endocervical specimens. Two of the ratios, 1544/1654 and 1240/1654, do not show significance between the epithelial layer and the basal layer, but the lamina propria layer can be differentiated from the other two layers. The difference in the 1338/1654 intensity ratio shows a higher ratio for the

epithelium layer, but a significant difference is not seen between the basal layer and lamina propria layer at this ratio.

The collagen and glycogen bands show the most noteworthy dissimilarities between the three layers at the 1080/1654 and 1025/1654 intensity peak ratios. In our preliminary data collection, all three layers can be discriminated spectrally at these regions, with the intensity ratio decreasing as the spectral collection occurs in deeper layers on the tissue specimen.

#### 2.4.1.2. Normal Ectocervix

We also evaluated the epithelium, basal, and lamina propria layers on normal ectocervical specimens to see if we could successfully differentiate. Variances again are depicted in the infrared spectra seen in Figure 2.70.

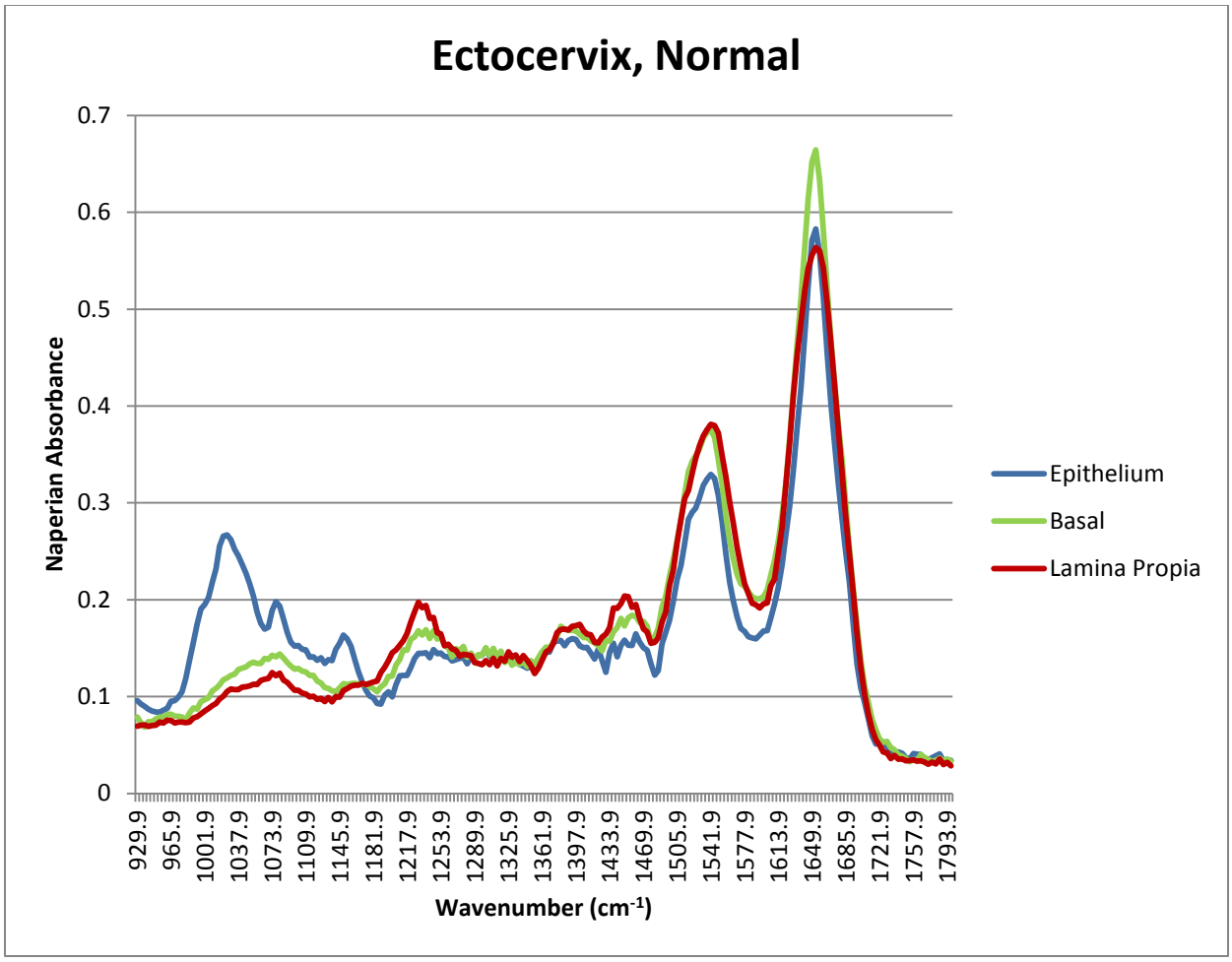



**Figure 2.70. IR spectra of average of normal ectocervical specimens, comparing epithelium, basal, and lamina propria layers**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 2.70 are shown in Figure 2.71. The plot shows the differences at peak intensity values of various bands in the spectra specific to the three layers being assessed.

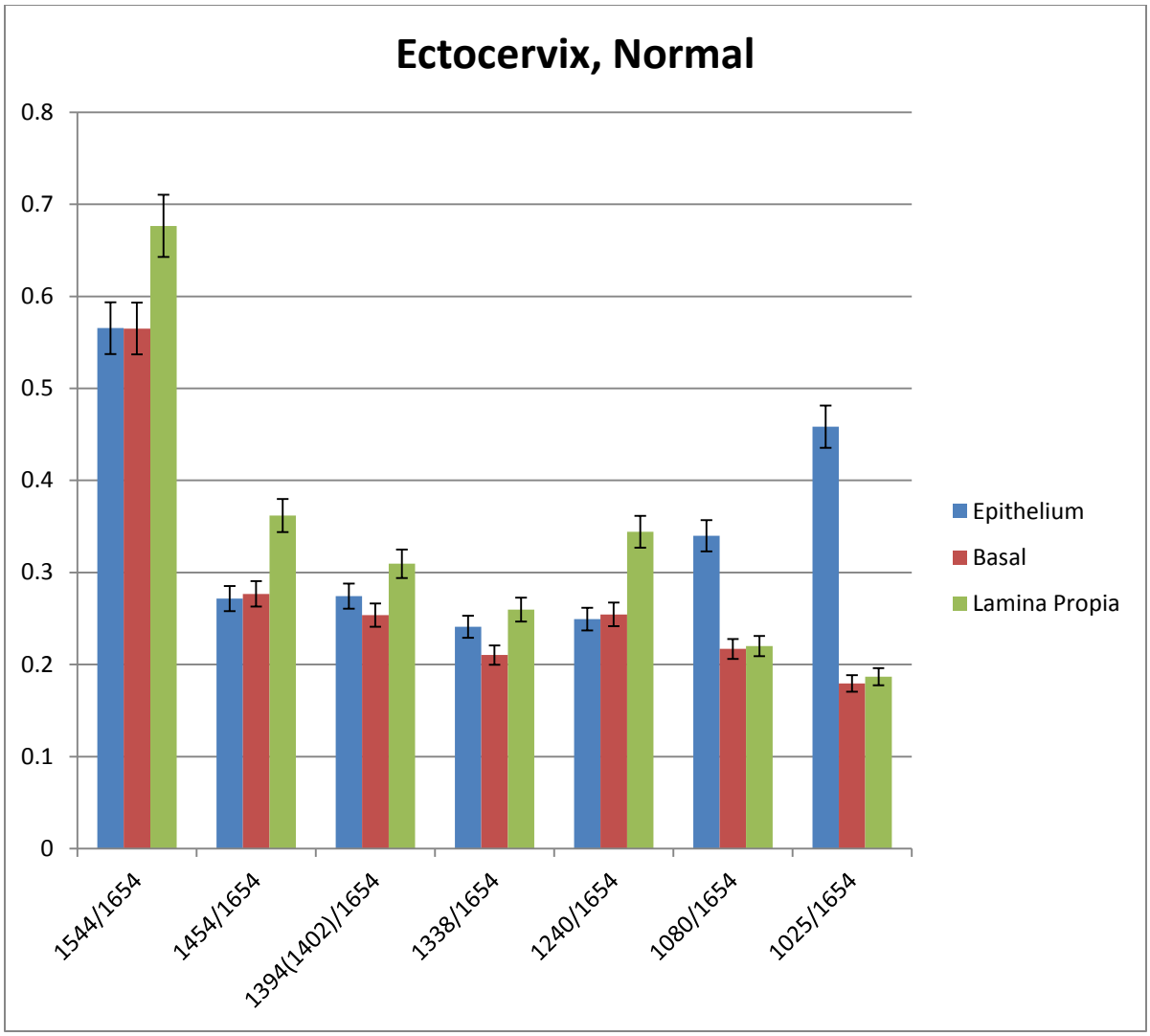



**Figure 2.71. Mean intensity ratio parameter of normal ectocervical specimens at different bands throughout the fingerprint region for epithelium, basal, and lamina propria layers**

All seven mean intensity ratio groupings show some statistical significance, although no one ratio allows discrimination among all three layers of the ectocervix. Four of the ratios, 1544/1654, 1454/1654, 1402/1654, and 1240/1654, do not show significant difference between the epithelial layer and the basal layer, but these ratios can be used to identify the lamina propria layer from the other two

layers. The difference in the 1338/1654 peak intensity ratio shows a lower ratio for the basal layer, but a significant difference is not seen between the epithelium layer and lamina propria layer at this ratio.

The collagen and glycogen bands show significant dissimilarity between the epithelium layer and the other two layers at the 1080/1654 and 1025/1654 intensity peak ratios. The basal layer and the lamina propria layer, however, cannot be differentiated between themselves at these two bands.

#### 2.4.1.3. Normal Comparison, Endocervix vs. Ectocervix

After comparing the epithelium, basal, and lamina propria cervical layers, we evaluated the differences strictly between the endocervical and the ectocervical specimens. The comparisons were broken into the three cervical layers, as depicted in Figure 2.72.

Comparing the infrared spectra of the endocervical epithelium and the ectocervical epithelium tissue specimens revealed no remarkable observable differences in the data. The endocervical samples appear to have slightly lower absorbance in the amide I and amide II protein bands. Again, declaring $p < 0.05$ as significant, the only intensity ratio parameters that bear any significance are at 1240/1654 and 1025/1654.

The basal spectral comparison is very similar to the epithelium. By observation, the spectra appear very similar, including the amide I and amide II

peaks. The only statistically significant intensity ratio parameter for basal tissue in the endocervix and ectocervix is the 1025/1654 ratio.

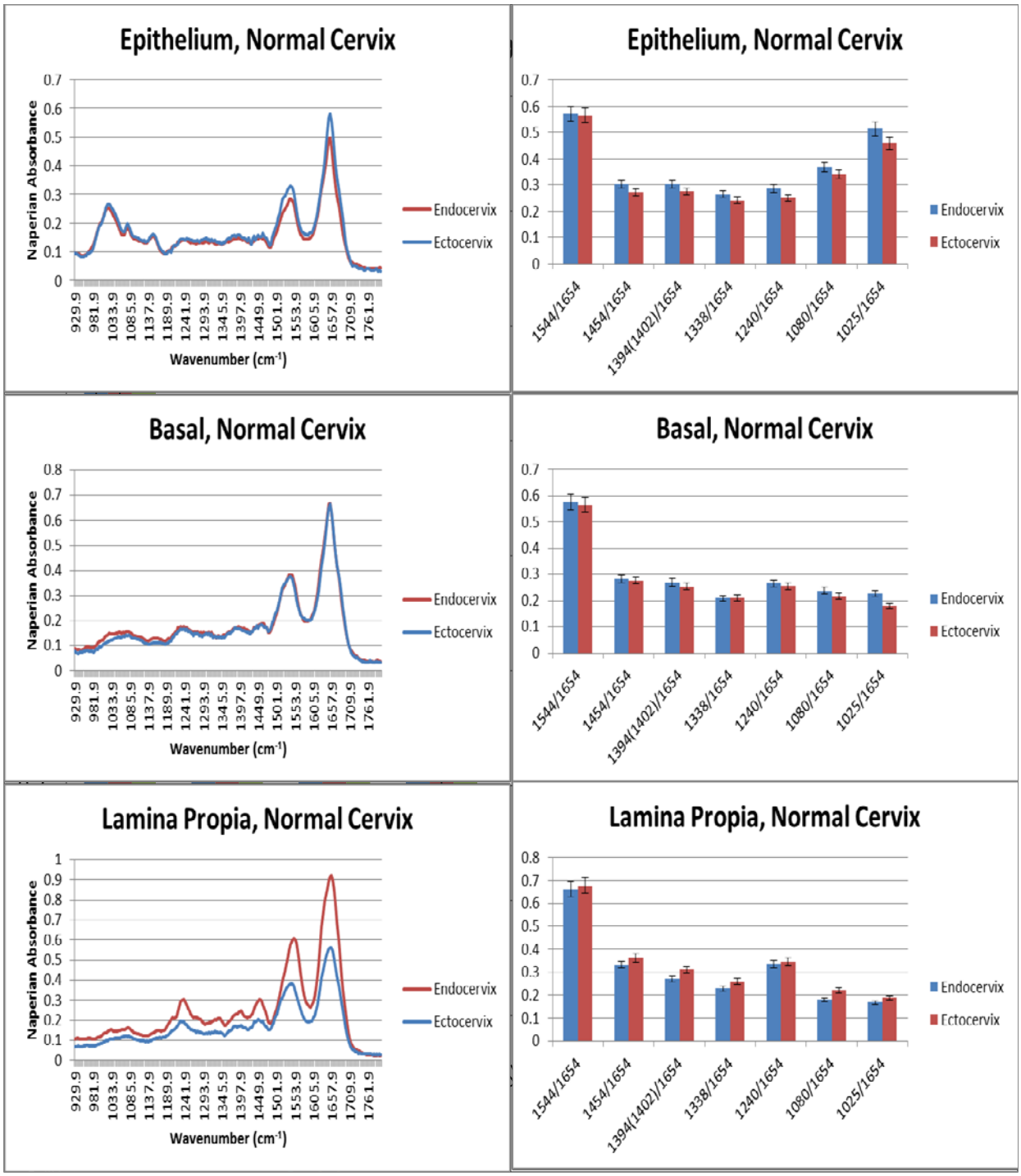



**Figure 2.72. Endocervix and ectocervix comparison. (Left Column) IR spectra of normal cervical specimens. (Right Column) Mean intensity ratio parameter of normal cervical specimens at different bands throughout the fingerprint region. ROWS: (Top) Epithelium, (Middle) basal, and (Bottom) Lamina propria comparison for endo/ecto cervix.**

The difference in the lamina propria absorbance spectra between the endocervix and ectocervix is clearly evident in Figure 2.72, although, based on the intensity ratio parameters, the scaling of the absorbance between the two is very similar throughout the entire spectrum. (This will be more evident in the area-under-the-curve discussions in Section 2.4.2.) The lamina propria has three intensity ratio parameters of significance at 1394/1654, 1338/1654, and 1080/1654.

#### 2.4.1.4. Malignant Squamous Cell Carcinoma and Glassy Cell Carcinoma

We next compared the mean spectral data between malignant squamous cell carcinoma specimens and the three specimens that were given the histological diagnosis by CHTN quality control personnel as "glassy cell carcinoma." The expectation had been very similar spectra between the two malignant sample types, but slight variations can be observed in Figure 2.73 between the squamous cell carcinoma and the glassy cell spectra. (This, too, will be more evident in the area-under-the-curve discussions in Section 2.4.2.) All spectral data displayed was collected from the samples' lamina propria layers.

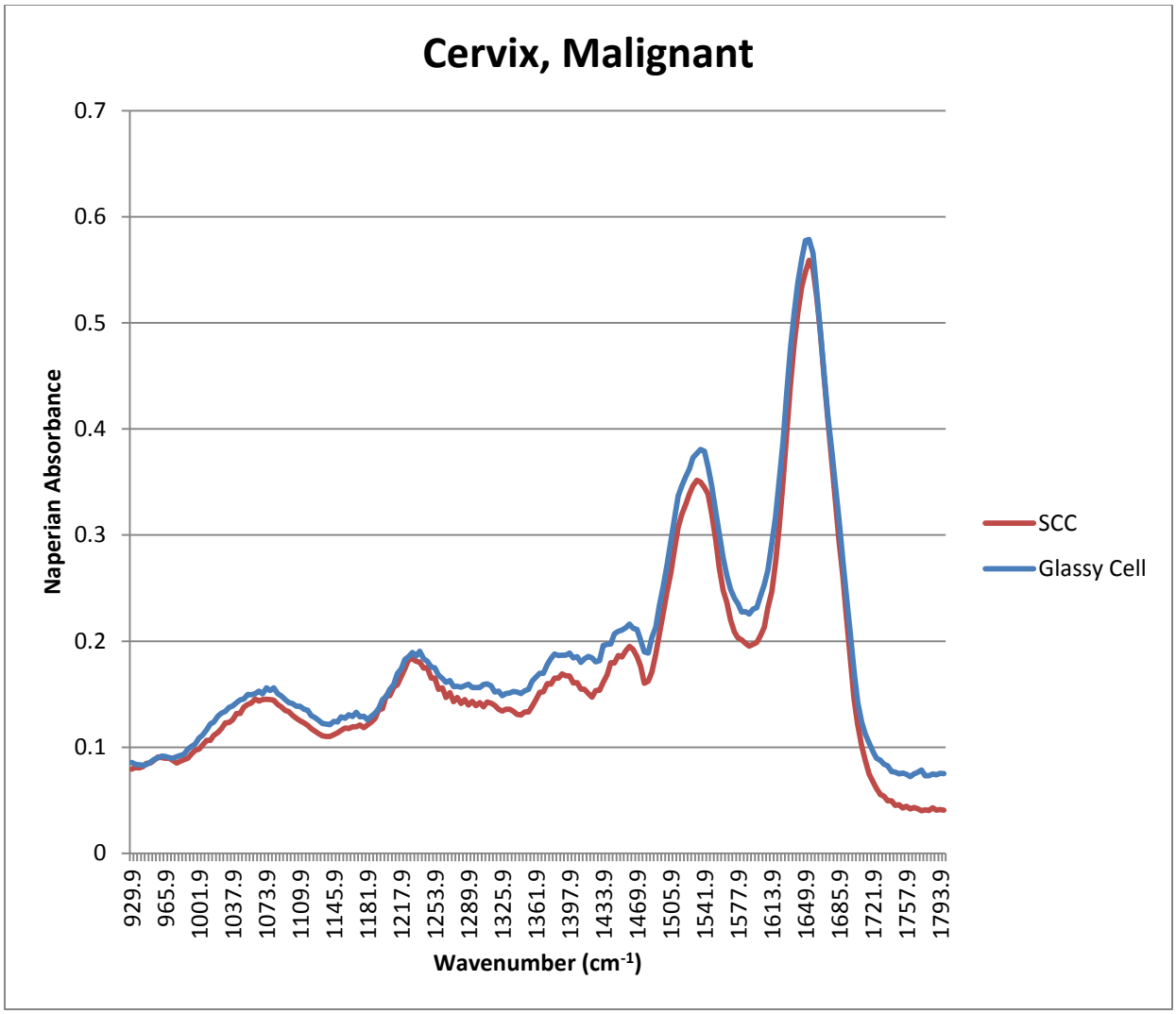



**Figure 2.73. IR spectra of malignant cervical specimens comparing squamous cell carcinoma to samples given glassy cell histological diagnosis**

The ratios of the mean intensity absorbance peaks are shown in Figure 2.74. The plot shows the differences at peak intensity values of various bands in the spectra specific to the two tissue types being assessed. No intensity peak ratios show statistical significance in differentiating between squamous cell carcinoma

and glassy cell carcinoma diagnosis, indicating any spectral difference in absorbance was scaled appropriately at each of the key absorbance peaks.

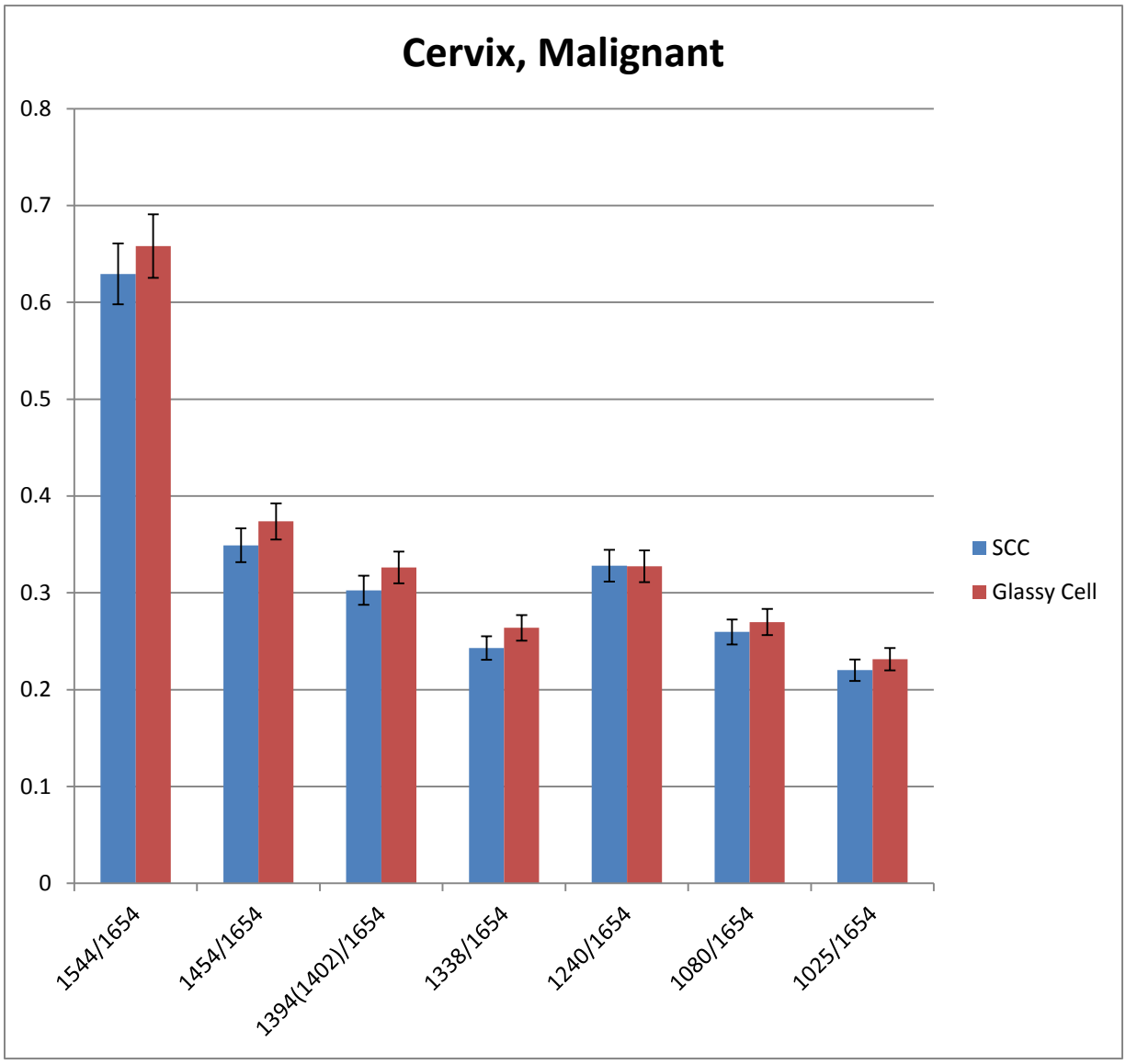



**Figure 2.74. Mean intensity ratio parameter of malignant cervical specimens at different bands throughout the fingerprint region for squamous cell carcinoma and glassy cell histological diagnosis**

### 2.4.1.5. Lamina Propria Comparison, Normal vs. Malignant

The final intensity ratio comparison we investigated was evaluating normal versus malignant cervical specimens. In order to strictly compare normal versus malignant tissue and not influence the results with other factors such as tissue layers, we only looked at the lamina propria layers in both the normal and malignant tissue samples because we were unable to discern the epithelium or basal layers in all our samples.

Clear variances are depicted in the absorbance intensity seen in Figure 2.75, which are the mean spectra of the malignant and normal tissue specimens. The average absorbance can be seen to be greater in the normal tissue throughout the fingerprint region with the exception of the collagen and glycogen bands around 1025 $cm^{-1}$ and 1080 $cm^{-1}$.

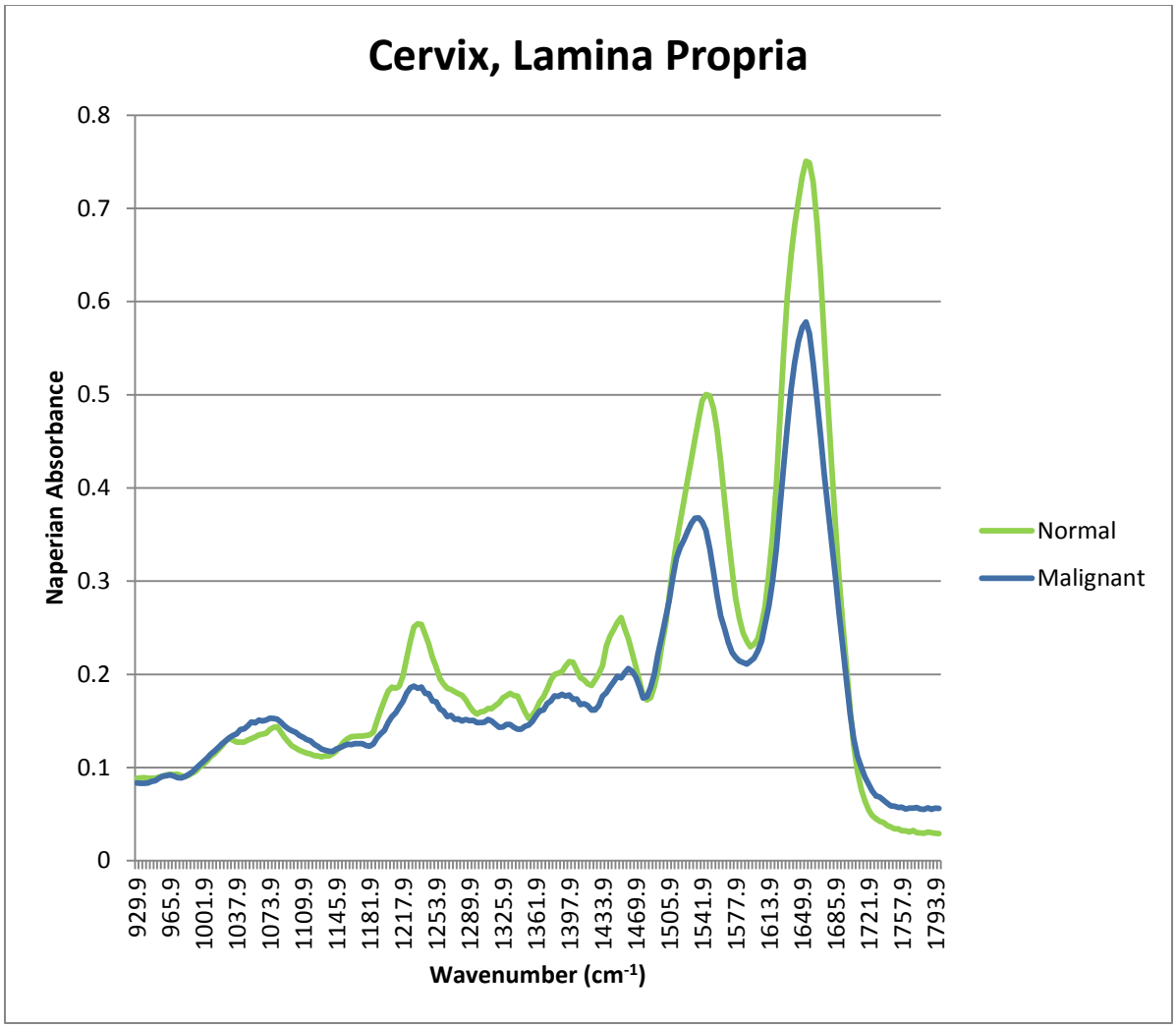



**Figure 2.75. IR spectra comparing average absorbance of normal and malignant cervical tissue**

The ratios of the primary mean intensity absorbance peaks seen in the spectra in Figure 2.75 are shown in Figure 2.76. The plot shows the differences at peak intensity ratio values of various bands in the spectra specific to the normal and malignant tissue samples.

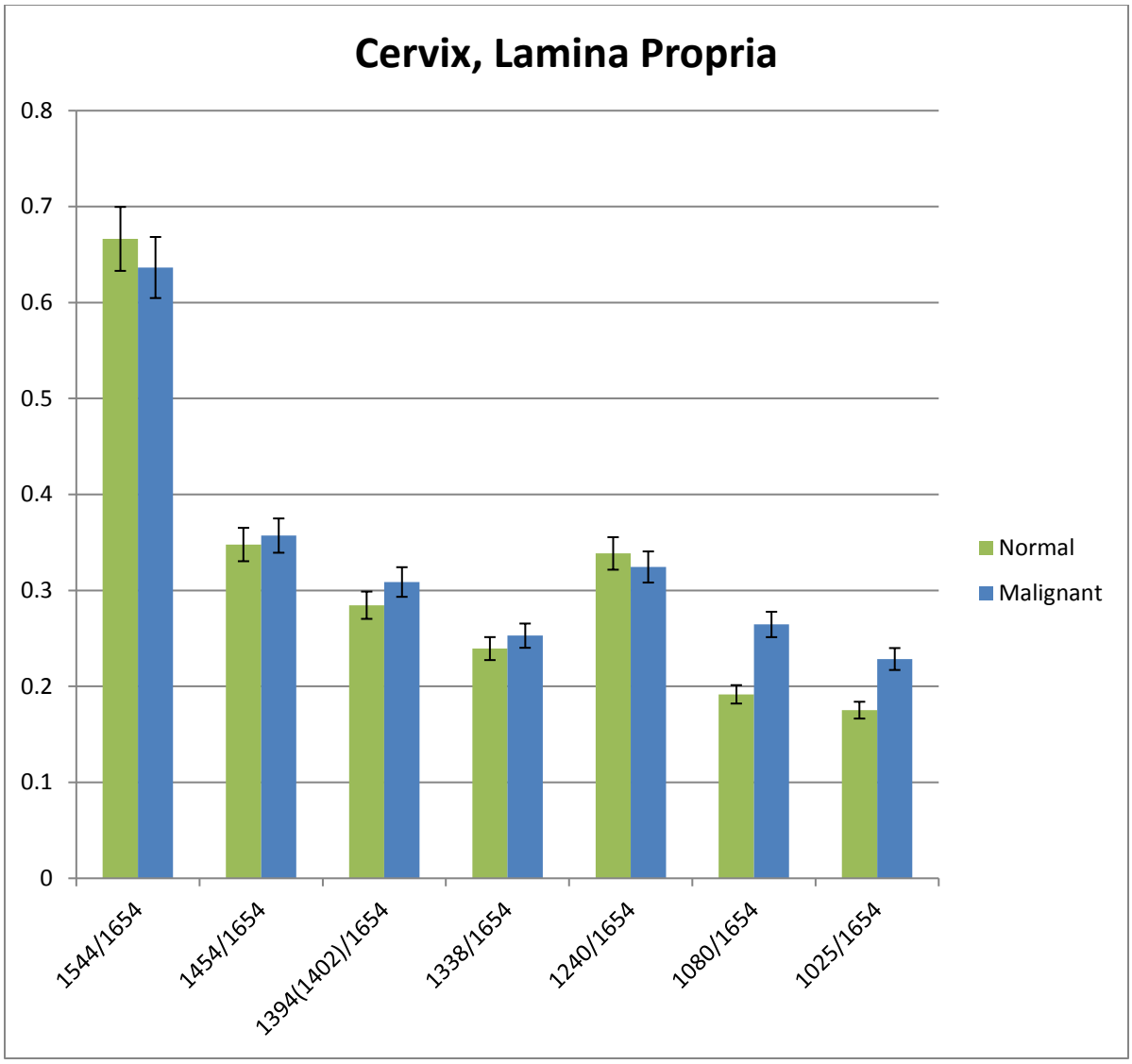



**Figure 2.76. Mean intensity ratio parameter of cervical specimens at different bands throughout the fingerprint region for normal and malignant specimens**

Only two of the seven intensity ratio groupings show statistically significant results for the lamina propria layer of the normal and malignant groups of specimens. The two ratios, 1080/1654 and 1025/1654, discriminate between the two tissue types being evaluated, indicating an increase of collagen and glycogen. In

fact, at 1025 $cm^{-1}$, the malignant spectrum did not have an absorbance peak. The glycogen at this point was saturated by the collagen signal of the 1080 $cm^{-1}$ band.

### 2.4.2. Area-Under-the-Curve Analysis

The area-under-the-curve (AUC) provides insight to absorbance intensity differences between spectra that are scaled similarly throughout the spectrum range, thus rendering the intensity ratio parameter analysis ineffective. The area under the curve was calculated using the trapezoid rule.

The AUC was calculated for all six spectra of the normal endocervical and ectocervical tissue spectra at the three layers we have been investigating. In Table 2.3, we can see the AUC differential between the six categories.

| | Area-Under-the-Curve | | |
|---|---|---|---|
| | Epithelium | Basal | Lamina Propria |
| Endocervix | 136.39 | 157.10 | 211.73 |
| Ectocervix | 147.07 | 150.37 | 146.45 |

**Table 2.3. Area-under-the-curve for the mean infrared spectra of normal epithelium, basal, and lamina propria layers of the endocervix and ectocervix samples.**

The endocervix shows substantial variance among all three layers being evaluated; there is significant difference among the AUCs to be able to distinguish the epithelium, basal, and lamina propria layers. Surprisingly, the ectocervix does

not have a significant difference among any of the three layers' AUCs. The layer variances are not significant enough to warrant solid classification of layers in the ectocervix.

Among the three layers being investigated, only the lamina propria shows a variance large enough to be significant in classifying between endocervix and ectocervix specimens. Both the epithelium and basal layers show too similar AUC values to be able to discriminate.

To note, the small sample size of specimens classified as endocervical or ectocervical specimens may have contributed to any surprising results. We did not expect to be able to vastly distinguish between ectocervix and endocervix samples, but we were anticipating being able to differentiate the epithelium, basal, and lamina propria layers in the ectocervix (just as we were able to in the endocervix spectra).

The AUC was also calculated for the squamous cell carcinoma specimens and those specimens diagnosed as glassy cell carcinoma. As seen in Table 2.4, we can determine the AUCs are just at the limit of being significant. Similar to the intensity peak ratio analysis, specimens diagnosed as glassy cell are too comparable to the standard squamous cell carcinoma specimens to be able to confidently differentiate.

| | Area-Under-the-Curve |
|---|---|
| | Lamina Propria |
| SCC | 149.43 |
| Glassy Cell | 165.57 |

**Table 2.4. Area-under-the-curve for the mean infrared spectra of malignant squamous cell carcinoma and glassy cell carcinoma.**

Finally, the AUC (shown in Table 2.5) was calculated from the mean IR spectra of all the normal and malignant cervical tissue samples. The variance between the two AUC values is significant enough to enable tissue classification using this method.

| | Area-Under-the-Curve |
|---|---|
| | Lamina Propria |
| Normal | 181.54 |
| Malignant | 158.51 |

**Table 2.5. Area-under-the-curve for the mean infrared spectra of normal and malignant cervical tissue.**

The values depicted in Table 2.5 coincide with the infrared spectra comparing normal and malignant tissue that were evaluated in the previous Section 2.4.1. Generally, throughout the infrared spectra in this fingerprint region, the

normal cervical tissue samples show greater absorbance than the malignant tissue samples.

### 2.4.3. Frequency Shift Analysis

Frequency shift analysis can only be tentatively performed with the current data set. Because the spectral data sets were collected at a 4 $cm^{-1}$ resolution, a shift of greater than 4 $cm^{-1}$ is needed to more fully exploit this analysis technique. However, some preliminary results with the current spectral data set can be defined.

The only absorbance peaks that show significant frequency shifts in the normal cervical tissue spectral data sets are the amide II protein band around 1544 $cm^{-1}$ and the two methyl protein bands around 1454 $cm^{-1}$ and 1402 $cm^{-1}$. The two methyl protein bands showed frequency shift within the normal cervical tissue samples, enabling differentiation between cervical layers. Our data set shows the 1454 $cm^{-1}$ band blue shifts in the basal layer by 8 $cm^{-1}$ to 1462 $cm^{-1}$ in both the endocervix and the ectocervix samples. In addition, the 1402 $cm^{-1}$ band red shifts in both the epithelium and basal layers by 8 $cm^{-1}$ to 1394 $cm^{-1}$ in both the endocervix and ectocervix. Finally, the ectocervix lamina propria red shifts from the endocervix lamina propria at the amide II protein band from 1550 $cm^{-1}$ to 1542 $cm^{-1}$. This was the only frequency shift that allowed us to differentiate between endocervical and ectocervical tissue samples.

Likewise, the absorbance peaks that show significant frequency shifts in the normal versus malignant cervical tissue spectral data sets are the same amide II protein band and two methyl protein bands; additionally, the collagen band at 1080 $cm^{-1}$ also revealed significant frequency shift. The absorbance peaks of the four bands mentioned clearly show shifts in the wavenumbers. The malignant tissue samples red shift 8 $cm^{-1}$ from the normal spectra at the amide II protein band at 1544 $cm^{-1}$, one of the methyl protein bands at 1402 $cm^{-1}$, and the collagen band at 1080 $cm^{-1}$. The malignant tissue also blue shifts by approximately 8 $cm^{-1}$ at the other methyl protein band at 1454 $cm^{-1}$. This data set is also the largest sample size, giving more confidence in the preliminary results.

## 2.5. Discussion / Conclusion

Mid-infrared spectroscopy has demonstrated an ability to provide biological information for cervical tissues that is inaccessible or more difficult to acquire with other techniques. Daylight Solutions' Spero instrument provided vast improvements over the standard FTIR spectrometer in the speed with which we collected the data and in the quality of our data collection.

In this chapter, we collected spectral data on a number of cervical tissue samples, both malignant and normal. Our focus on the normal tissue was the epithelium layer, the basal layer, and the lamina propria layer; our focus on the malignant tissue was the densest portion of the tumor, typically in the lamina propria layer since the epithelium and the basal layers were indistinguishable. We

were able to focus and pinpoint our data collection techniques to exact locations on the sample by the following: (1) obtaining H & E stains of an interior excised sample of each specimen and locating areas of interest, (2) taking an adjacent excised sample with which to collect spectral data on Spero, and (3) utilizing the mapping feature on Spero to match the location of interest previously defined on the H & E stain. The standard FTIR instrument would not have been able to collect spectral data strictly from the epithelium or basal layers because of the inability to pinpoint data collection sites.

We collected numerous images and spectra from each cervical specimen; some of the examples of these data assemblages are shown in this chapter and in Appendix A and Appendix B. Because such a wide variety of histology diagnostic information was included with the tissue specimens, we focused our efforts on comparing epithelium, basal, and lamina propria layers on normal cervical samples; normal endocervical and ectocervical samples; malignant samples diagnosed as squamous cell carcinoma and glassy cell carcinoma; and normal and malignant cervical samples (lamina propria layer). For the initial analysis using Daylight Solutions' Spero microscope, we used the following three analysis techniques: intensity ratio parameter analysis, area-under-the-curve analysis, and frequency shift analysis.

Our preliminary results enabled us to differentiate between all three cervical layers by analyzing the spectral data sets using intensity ratio parameter analysis. This initial analysis proved effective with our data, whether evaluating endocervix

or ectocervix specimens. Calculating the area-under-the-curve for each spectrum proved less effective. While we were able to identify the three cervical layers in the endocervix utilizing the preliminary results, we were not able to significantly differentiate the three layers in the ectocervix. This appears to be because the amide I and amide II peaks (the two most dominant bands in the fingerprint region) have similar absorbance in the ectocervix between the basal and lamina propria layers, and although the epithelium layer still has higher absorbance in the glycogen and collagen bands, the amide I and amide II absorbance is slightly lower than the other two layers. The endocervix layers have clear distinctions in the amide I and amide II bands.

Although we were able to discern the endocervix from the ectocervix using the preliminary intensity ratio parameter analysis while still fulfilling the $p < 0.05$ criteria, the variation was not nearly as significant as with the three epithelium, basal, and lamina propria cervical tissue layers. The AUC proved slightly more effective. Among the three layers being investigated, only the lamina propria preliminary results show variances large enough to be significant in classifying between endocervix and ectocervix specimens. After observing the spectra of the three cervical layers, these results were expected.

The initial intensity peak ratio analysis did not prove effective in differentiating between squamous cell carcinoma and glassy cell carcinoma, although this result was expected since both specimen types are malignant tumors, the main differences being the “glassy-like” appearance of the cytoplasm and the

aggressive nature of the tumor. Although the AUC preliminary analysis did show some variation between these two types of malignant samples, the variation was close to the criteria limitation. Due in part to the fact that we were only able to acquire three glassy cell carcinoma specimens, we do not have high confidence in these results.

Finally, we were able to successfully identify normal and malignant specimens with our preliminary results. Observably, comparing the spectra against each other, we could see visual qualitative differences between the two tissue types. We analyzed the spectral fingerprint using the intensity ratio parameter analysis, and the initial results show that we may be able to distinguish with significance the two specimen types, mainly due to changes in the glycogen and collagen. We also calculated the areas-under-the-curve for these two spectra and found large enough variances in the initial values that may allow us to discriminate between normal and malignant cervical samples.

We did preliminarily look at frequency analysis in all of our comparisons as well. Although this technique shows promise as being a clear way to differentiate among cervical layers and between malignant and normal cervical specimens, more experiments with higher resolution need to occur. Specifically, because one of the bands with a large frequency shift is the highly absorbent amide II peak, this is a technique that could prove quite useful in separating the normal and malignant samples. Often, while researchers are using FTIR spectrometers, 4 $cm^{-1}$ seems to be a standard resolution during data collection. Any higher resolution would extend

the already long data collection process. Utilizing Spero's tools could specify precise wavenumbers and/or bands while increasing the resolution of data points to 1 $cm^{-1}$ or less; the data collection time would be on the order of minutes instead of several hours.

The normal cervical experiments specifically involving the three cervical layers and the endo/ectocervix had small sample sizes, which could contribute to any surprising results. Although we had twenty normal cervical specimens, only eight were identified as endocervix and seven as ectocervix. Some of these specimens were removed because a clear epithelium edge was not defined on the sample. Some samples were also removed due to histology diagnostic comments indicating they were a benign tumor; although these specimens are not malignant, they still contained a tumor and were thus removed so as to not taint the preliminary results. (The benign normal specimens were also removed from all other results.)

More experiments with a larger sample size need to be conducted in order to substantiate these results, which should also include specific specimens with defined histology diagnostic characteristics. Multivariate analysis techniques could be considered in order to tease out more subtle information available in the spectra. Our work with the Spero quantum cascade laser microscope demonstrates great advantages over the standard FTIR spectrometer and is the first step in finally making great strides in the mid-infrared cancer detection field that has been

stagnant for so many years. With further exploitation of Spero's advantages, the speed of data collection and analysis can be improved ten-fold.

# Chapter 3

# Mid-IR Spectral Investigation of Normal and Malignant Breast Tissue Samples

Cancer is the second leading cause of death in the United States, and 8.5% of all Americans have been diagnosed with cancer, which, according to the Center for Disease Control and Prevention, accounts for 23% of deaths in the United States. The survival rate for people diagnosed with cancer greatly increases with early detection.[61-63]

The second most common form of cancer worldwide is breast cancer, and it is the most common form of cancer in women. According to the World Cancer Research Fund International, 1.7 million people were newly diagnosed with breast cancer in 2012, which represents ⅛ of total cancer cases and ¼ of all cancers diagnosed in women. The most alarming global statistic is the increased rate of breast cancer incidences and mortality cases. In only the last six years, breast

cancer diagnoses have increased by 20%, and breast cancer mortality rates have increased by 14%.[61-63]

In the United States, breast cancer is the second leading cause of cancer deaths in women; over 40,000 women will die from breast cancer this year. If the breast cancer can be diagnosed and identified before spreading to lymph nodes, the five-year survival rate is 98.5%. However, if the cancer is not diagnosed in time and spreads to regional lymph nodes, survival rate drops to 84%; if it spreads to distant lymph nodes, the survival rate is only 24%.[61-63]

Breast cancer is related to hormonal levels; however, diagnosis can be difficult because hormones are inconsistent from person to person, and an individual's hormones change from year to year, especially pre-menopausal and post-menopausal. Environmental factors also play a role in a person's hormonal levels. Because of these inconsistencies, alternative forms of diagnosis are continually being investigated.

Quick, easy, and accurate detection in breast cancer and precancer can reduce medical costs, time between biopsy and patient results and treatment, and assist with automated prescreening of tissue readings. In this study, we applied Daylight Solution's quantum cascade laser-based microscope (Spero) to investigate mid-infrared characteristics of normal and malignant breast tissue. First, sample thickness was evaluated by sending various thickness samples to Daylight Solutions for evaluation. Second, samples fixed in paraffin were evaluated as is and de-

paraffinized. Finally, the normal and malignant samples were prepared and evaluated utilizing the unique capabilities of Spero.

## 3.1. Materials

### 3.1.1. Spero

Refer to Section 2.1.1 Spero for instrumentation description and specifications used in the experiment and research that are described in this chapter.

### 3.1.2. Tissue Samples[52]

We acquired all our breast tissue samples, both normal and malignant, from the Cooperative Human Tissue Network (CHTN) based out of the University of Alabama Birmingham. Table 3.1 and Table 3.2 show additional histological information provided by CHTN with each breast sample. The normal tissue is only categorized by ethnicity of the patient; location of the specimen was not included, and no other relevant statistical information was provided. Because CHTN classified the ethnicity of all the samples as Black, White, Asian, or Hispanic, we used these same labels throughout our documentation.

The malignant tissue specimens did include more statistical information, such as location on the breast (ductal or lobular). The malignant breast tissue is also categorized by ethnicity of the patient from whom the specimen was removed.

Finally, receptor information was included for approximately half the patients, specifically the two hormonal receptors estrogen and progesterone and the protein receptor HER2.

| | | |
|---|---|---|
| Ethnicity | White | 17 |
| | Black | 3 |
| Total Breast Normal Samples: | | 20 |

**Table 3.1. Normal breast samples.**

| | | |
|---|---|---|
| Breast Location | Ductal | 9 |
| | Lobular | 5 |
| | Phylloides | 1 |
| Receptors | ER/PR+, Her2- | 4 |
| | ER+, PR/Her2- | 2 |
| | ER/PR/Her2- (triple neg) | 1 |
| | Unknown | 8 |
| Ethnicity | White | 10 |
| | Black | 4 |
| | Asian | 1 |
| Total Breast Cancer Samples: | | 15 |

**Table 3.2. Malignant breast samples.**

The tissue samples included sparse diagnostics and comments received from the medical staff and principle investigators cooperating with CHTN. The specimens went through an initial diagnostic examination by a trained pathologist, which were then verified through a frozen section and evaluation of permanent histopathology. A total of twenty normal breast specimens and fifteen malignant breast specimens were received from CHTN.

### 3.1.3. Optical Slides and Windows

Refer to Section 2.1.3 Optical Slides and Windows for material description and specifications used in the experiment and research described in this chapter.

### 3.1.4. Biology of Breast Tissue Specimens[64-66]

The breast is mainly composed of tissue, fat, glands, and ducts; the majority of breast cancer occurs in either the ducts or the glands. The lobules are the milk-producing glands located inside the breast. The ductal tissue is the ducts that transport the milk from the lobular glands to the nipple. The lobules are located at the inner-most end of the milk ducts.

Ductal carcinoma in situ is cancer that starts and remains in the ductal tissue (refer to Figure 3.1)[64]. It is the most common form of non-invasive cancer. Although this cancer is not fatal, it can lead to invasive ductal carcinoma or indicate higher susceptibility to other invasive forms of cancer.

Invasive ductal carcinoma is the most common form of breast cancer, accounting for approximately 80% of all breast cancer cases. Invasive ductal carcinoma has spread from the milk ducts to surrounding breast tissues. Most cases of invasive ductal carcinoma occur in women older than 55, although this cancer can affect women of any age and can also occur in men.

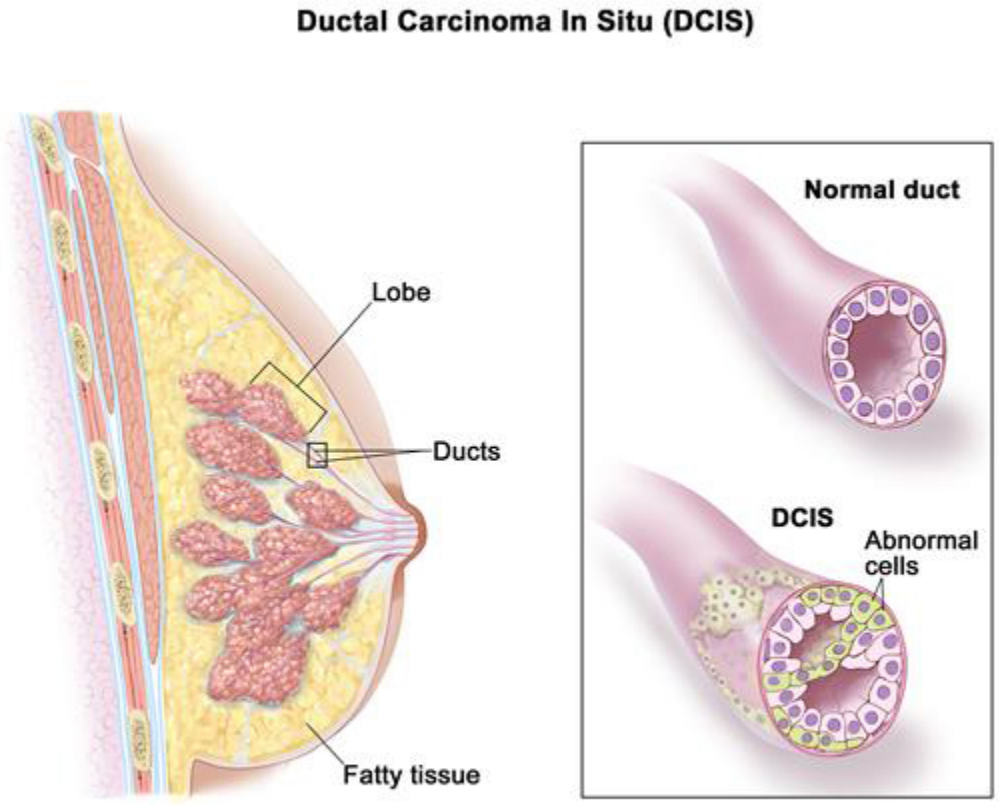



**Figure 3.1. Ductal carcinoma in situ and duct location in the breast.[64]**

Lobular carcinoma in situ, also referred to as lobular neoplasia in situ, denotes to abnormal cell growth in the lobular glands of the breast (refer to Figure 3.2)[64]. Typically, a person diagnosed with lobular carcinoma in situ will have more than one lobule affected. Although not actually a cancer, a diagnosis of lobular carcinoma in situ is an indicator of higher susceptibility for breast cancer at a later time. Over 90% of lobular carcinoma in situ diagnosis occurs in pre-menopausal women between the age of 40 and 50. It can be difficult to diagnose due to the fact

that no external physical symptoms have been identified, and it is typically found as a result of unrelated testing. Lobular carcinoma in situ occurs very rarely in men.

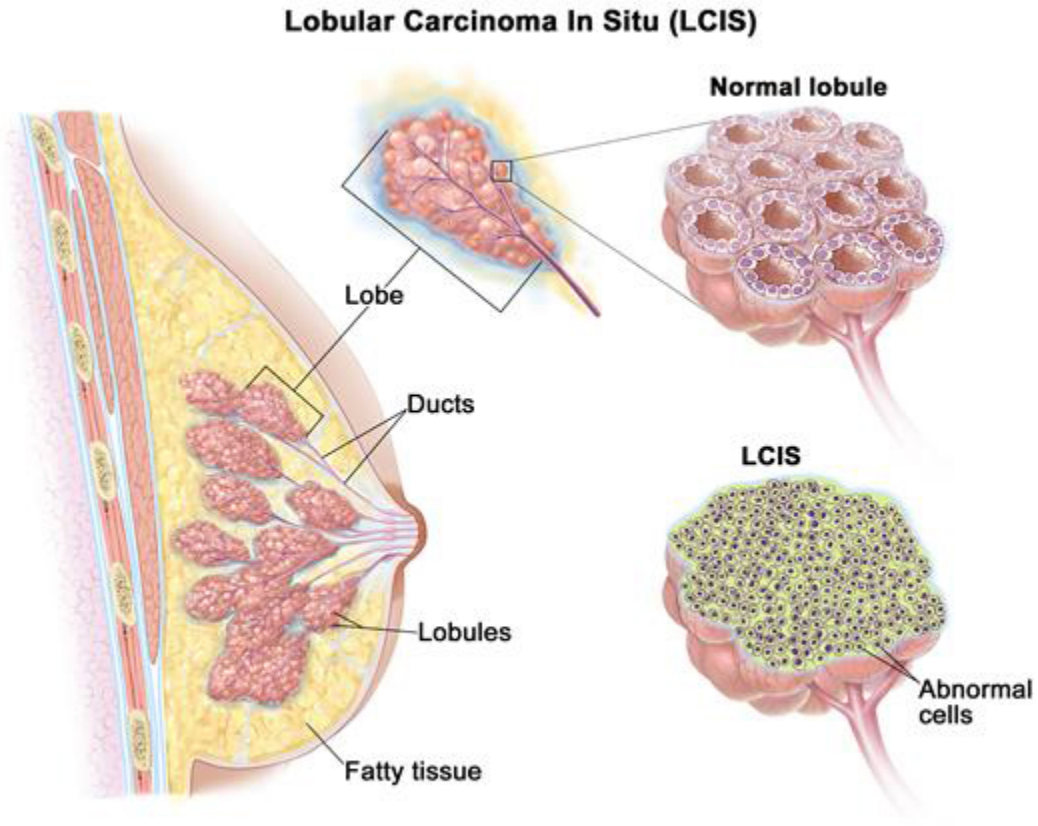



**Figure 3.2. Lobular carcinoma in situ and lobule location in the breast.**[64]

Invasive lobular carcinoma is the second most common type of breast cancer, accounting for approximately 10% of all breast cancer incidents. Invasive lobular carcinoma has spread from the milk-producing lobule glands to the surrounding

breast tissues. Typically, invasive lobular carcinoma occurs in women later in life than invasive ductal carcinoma, although it can occur in women of any age.

## 3.2. Methods

### 3.2.1. Sample Preparation

The breast tissue samples arrived frozen from CHTN in dry ice. Baylor College of Medicine (BCM) Pathology and Core Lab was utilized to prepare the tissue samples. The tissue specimens were all fixed and embedded in paraffin.

Daylight Solutions has evaluated various sample thicknesses to use on Spero using other tissue types. They have determined that the optimal thickness for transmission data collection is between 5 and 7 μm. They also concluded the optimal thickness of the CaF2 windows to be 2 mm and provided a holder for ICL's 25 mm diameter disc.

BCM's Pathology and Core Lab fixed each specimen in formalin; the specimens were then each embedded into a block of paraffin. Two tissue samples, approximately 5-7 μm thick, were excised from each specimen; one sample was placed on a standard glass slide, and one sample was placed on a 2 mm thick calcium fluoride (CaF2) window from ICL. The samples were then de-paraffinized to prevent artifacts in the spectra.

Samples on the standard glass slides were stained with hematoxylin and eosin (H & E stain), as seen in Figure 3.3. These were used to identify regions of

interest on the CaF2 samples (see Figure 3.4), which were used for transmission data collection on Spero. All samples were then stored in a desiccator while not in use.

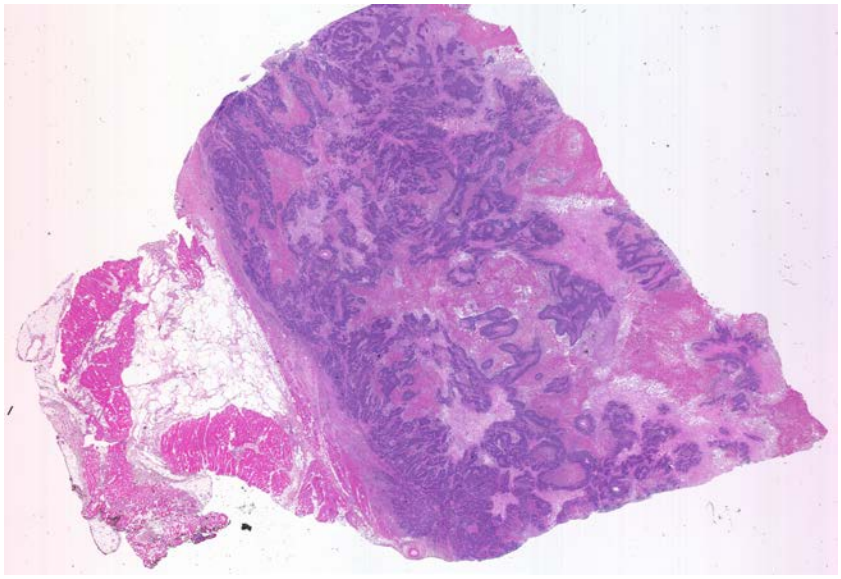


**Figure 3.3. Example of H&E stain, malignant breast specimen 10-10-A177m.**

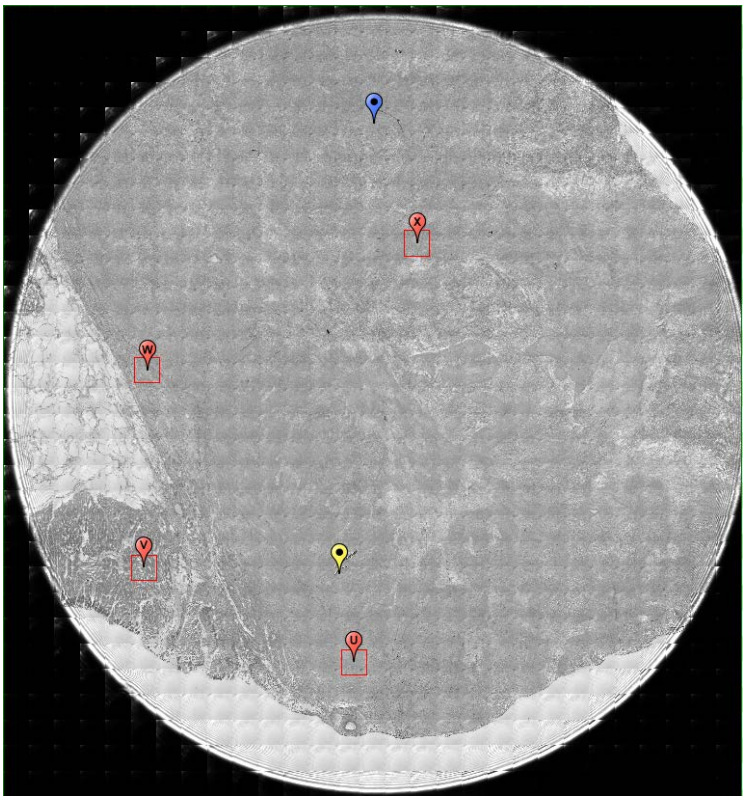



**Figure 3.4. Example of Spero image, malignant breast specimen 10-10-A177m.**

### 3.2.2. Data Acquisition Process

The Spero microscope was set up on a floating optical table to absorb environmental vibrations and other noise interference. A constant flow of nitrogen gas at approximately 5 cc/second was required to purge the sample stage enclosure of water vapor and other contaminants. Initially we obtained nitrogen tanks to perform the purge, but the tanks only lasted 1-2 days of continuous 5 cc/second flow. We then changed to the facility's nitrogen ports available in the lab by running approximately 80 feet of hosing to the nearest nitrogen port.

Prior to any use, we ensured the microscope's lasers had sufficiently warmed up by having the instrument on for at least an hour. The back of each 25 mm diameter CaF2 window was gently cleaned with optical tissue and an ethanol solution to remove dust particles. The sample was placed on a holder provided by Daylight Solutions and positioned on the stage. The stage cover was then closed and the system was allowed to purge with nitrogen for at least five minutes prior to any data or background collection. A new background was taken on a clean portion of the same slide the sample was on prior to any data collection of a new sample; a new background was taken at a minimum every three data points of that sample or every 45 minutes.

Each Spero image was attained using the high-magnification infrared objective. The tissue sample was oriented similarly to the H&E stained image, as seen in Figure 3.3 and Figure 3.4. Each sample on a CaF2 window was compared to the H&E stain to identify regions of interest to collect data points. The focus of data collection on the normal specimens was on more dense areas of the breast (non-fatty), if possible. The focus of the malignant specimens was on denser cell areas indicating the tumor location.

### 3.2.3. Peaks Corresponding to Biology

Refer to Section 2.2.3 Peaks Corresponding to Biology for specific wavenumbers and ranges of wavenumbers in the mid-infrared spectrum that

correspond to unique biological characteristics that are used in the experiment and research described in this chapter.

## 3.3. Results

Mid-IR spectral areas of the highest absorption in the breast samples tend to be the two protein Amide peaks (1650 $cm^{-1}$ and 1545 $cm^{-1}$), along with phosphates (asymmetrical stretching at 1240 $cm^{-1}$ and symmetrical stretching at 1080 $cm^{-1}$). Smaller peaks also appear at methyl bonds of protein at 1400 and 1450 $cm^{-1}$.

### 3.3.1. Normal Breast Data Acquisition

Twenty normal breast specimens were utilized for this experiment. Samples from each specimen were placed on glass slides and CaF2 windows; the samples on the glass slides were stained with H&E. Data for each sample on the CaF2 windows were collected on Spero and evaluated in this experiment; the locations of the data collection points were first identified on the samples stained with H&E. Each sample had multiple points of data collection, with the emphasis placed on areas of denser tissue since the fat cells tended to wash out in the paraffin / de-paraffin process. A few key normal breast specimen examples are listed in this section. Refer to Appendix C for a complete listing of normal breast specimens, including images from the H&E-stained slides, the images obtained through Spero's high-magnification objective, and the spectra from each data point collected.

#### 3.3.1.1. Specimen 11-08-A085c

Specimen 11-08-A085c is a 4.70 gram sample from a 31-year-old white female. Figure 3.5 is an image of a sample of the specimen stained with H&E on a glass slide. The pink and purple areas indicate denser areas of the sample; the fat cell regions can be faintly seen in areas on the sample by the cellular membranes left from the fat cells washing out during the preparation process.

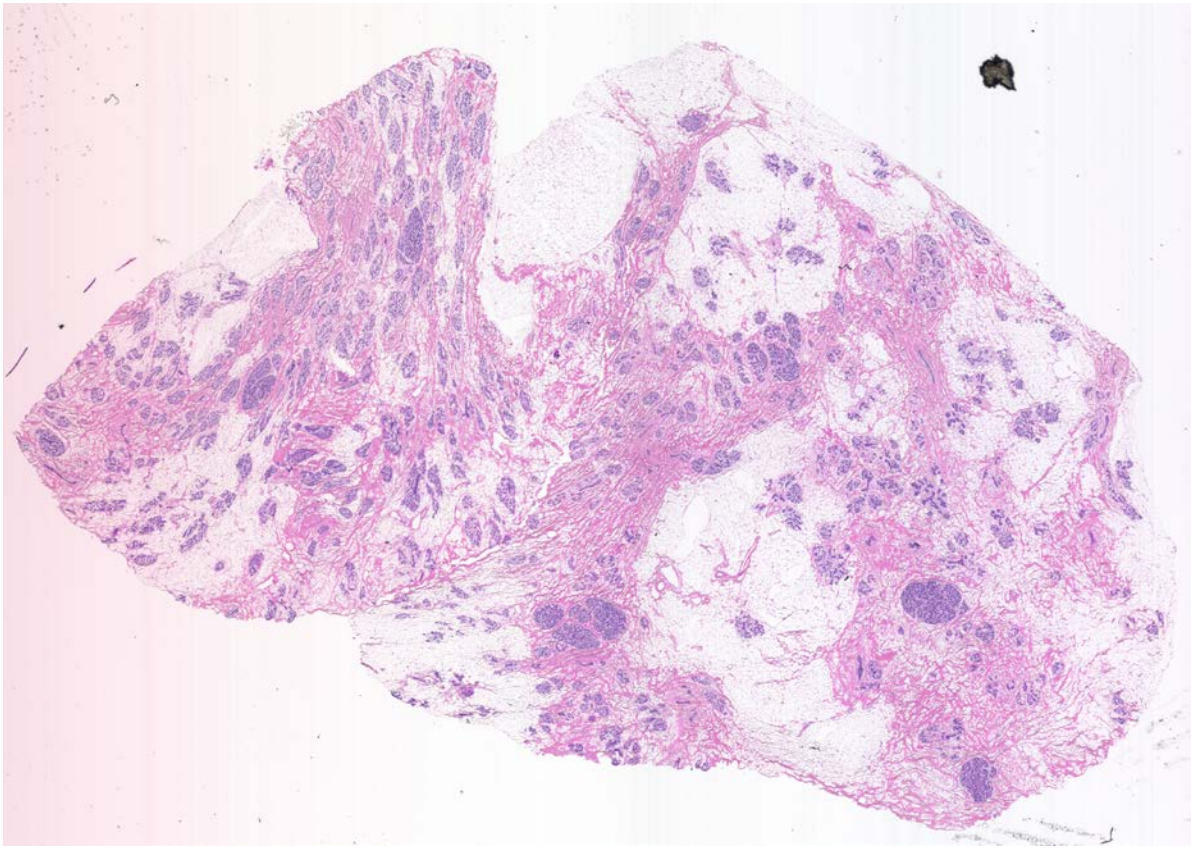


**Figure 3.5. Specimen 11-08-A085c, H&E stain, normal breast tissue.**

Figure 3.6 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags J and K depict the data collection points on Sample 11-08-A085c, focusing on the denser tissue areas. Flag L is a data collection point taken on fat cells that were washed out during the paraffin / deparaffin process.

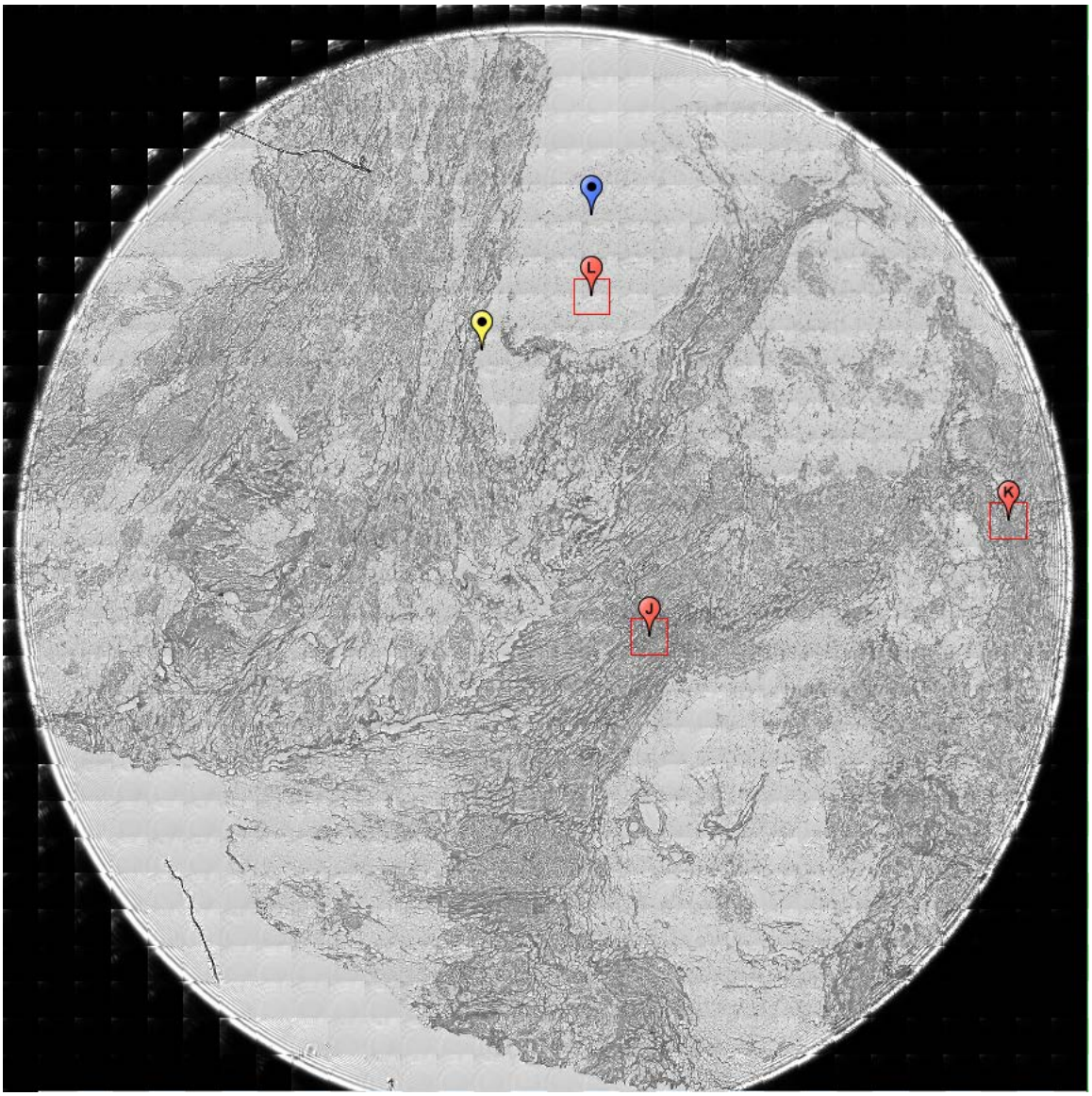



**Figure 3.6. Specimen 11-08-A085c, normal right breast tissue.**

#### 3.3.1.1.1. Sample 11-08-A085c, Point J

A 650 x 650 μm field-of-view image taken of the specimen at Point J (see Figure 3.7) was obtained from Spero's high-magnification objective. Only one

region of interest was needed to capture the spectral information at this point on the sample.

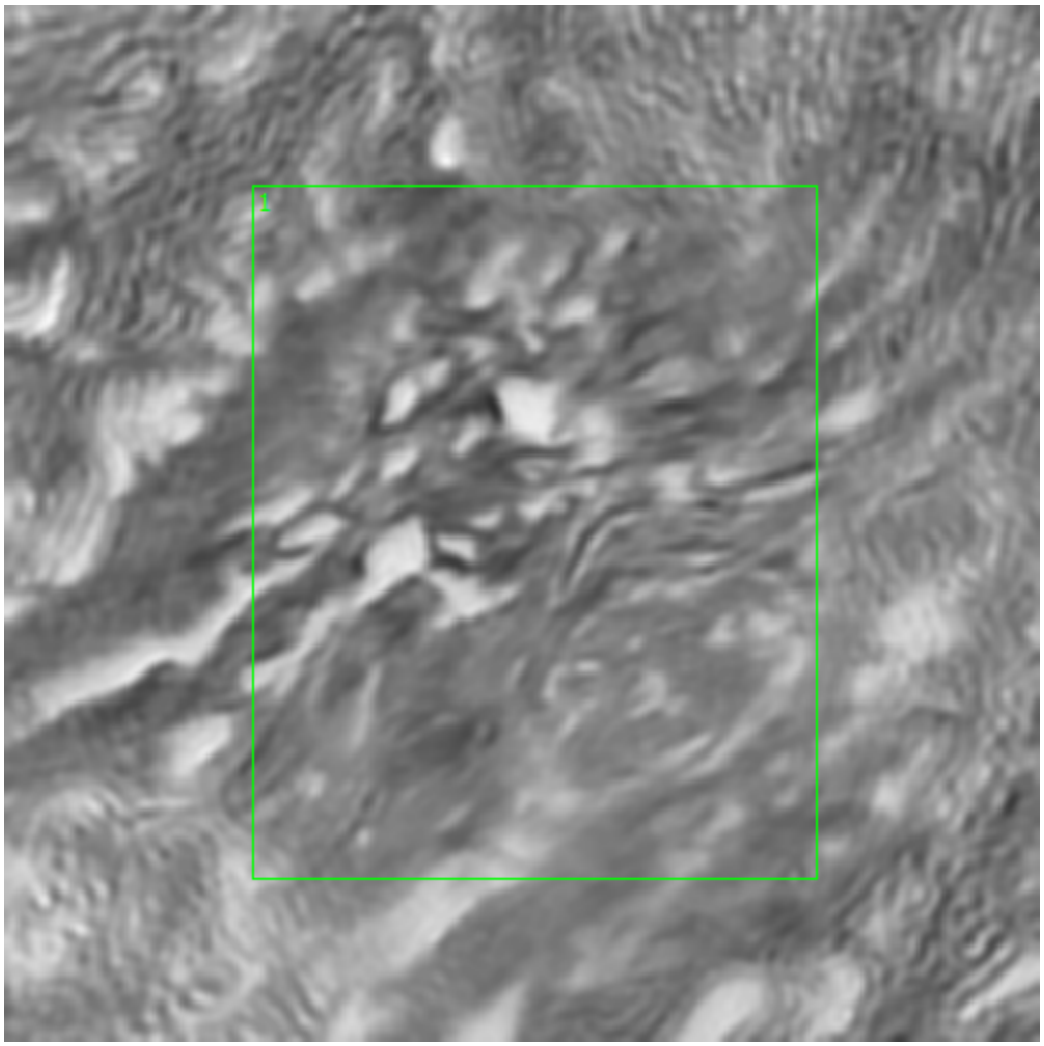



**Figure 3.7. Specimen 11-08-A085c, Point J Regions of Interest**
**Region 1: normal right breast tissue**

The spectrum from the defined region of interest on Point J is depicted in Figure 3.8.

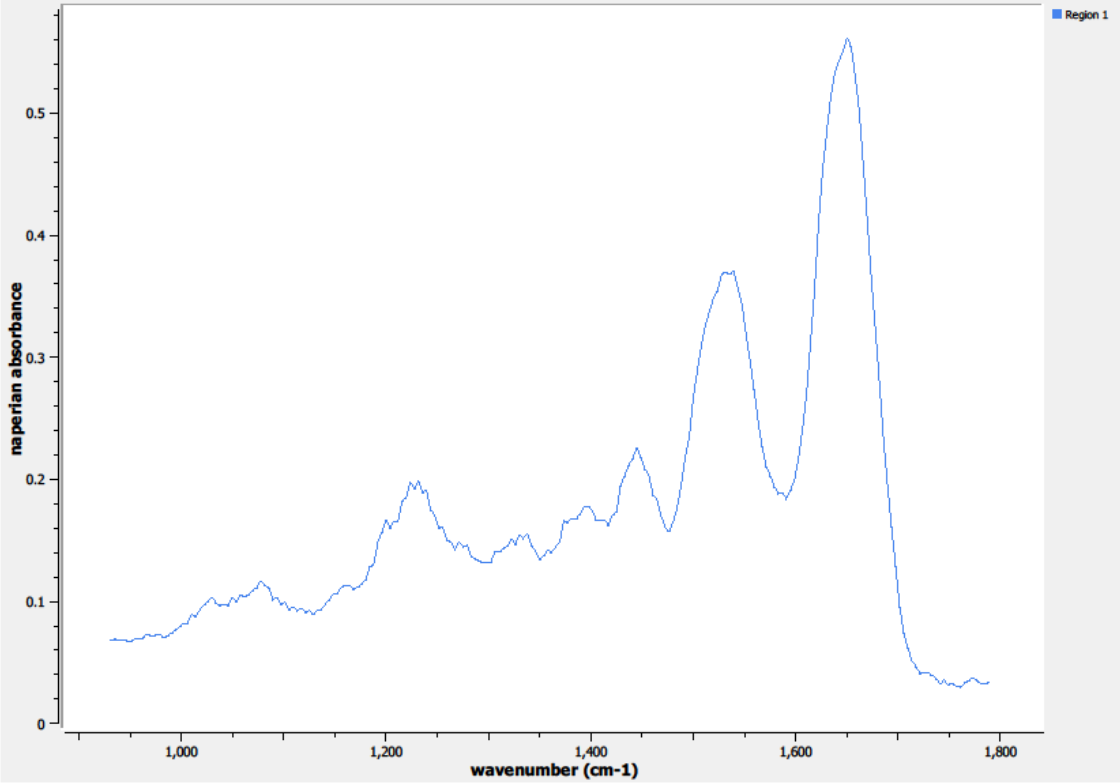



**Figure 3.8. Specimen 11-08-A085c, Point J Spectra**

#### 3.3.1.1.2. Sample 11-08-A085c, Point K

The high-magnification 650 x 650 µm field-of-view image was also taken at Point K (see Figure 3.9) on Sample 11-08-A085c. Region 1 is a general region of interest over the center of the field-of-view; Region 2 is defining a more dense tissue area for spectral data collection.

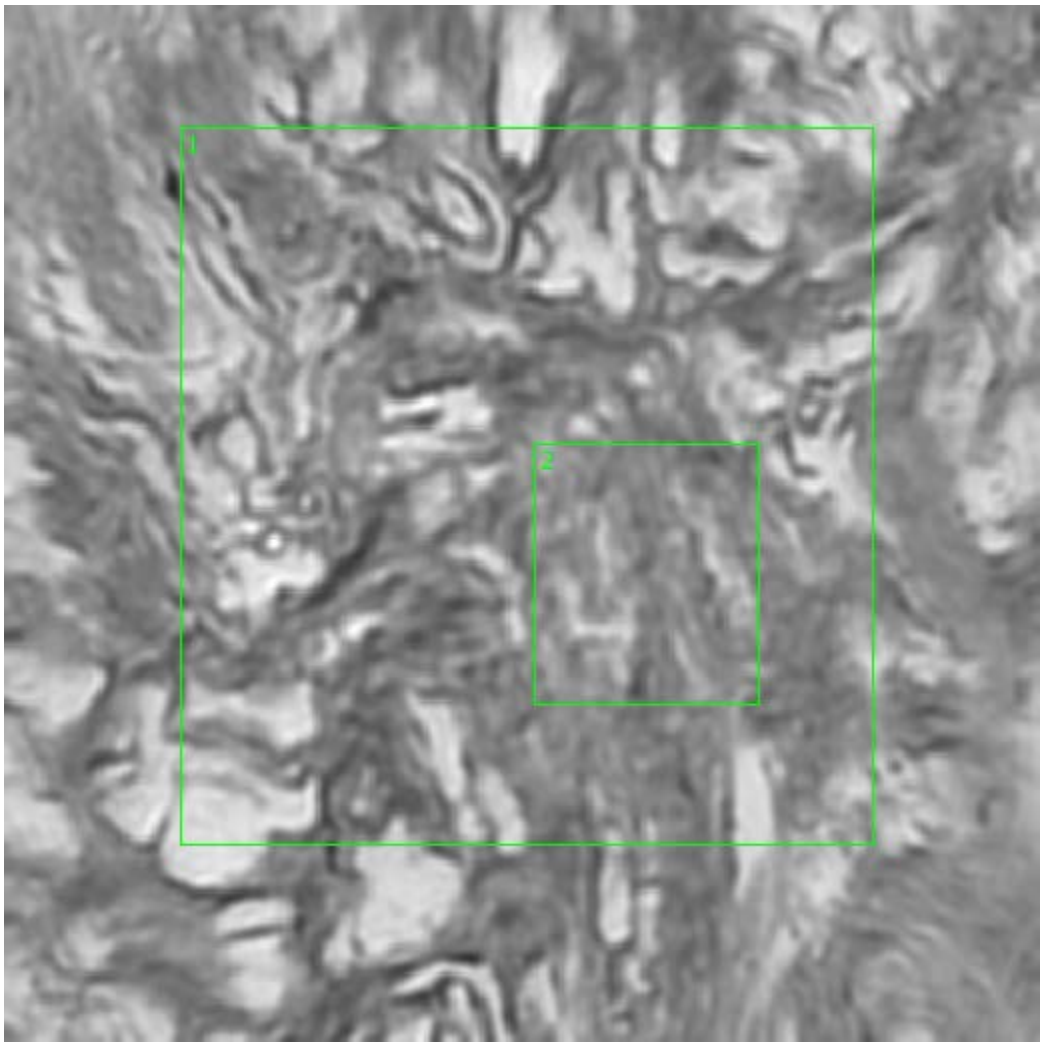



**Figure 3.9. Specimen 11-08-A085c, Point K Regions of Interest**
**Region 1: normal right breast tissue**
**Region 2: concentrated region, normal right breast tissue**

The spectra from the two defined regions of interest on Point K are depicted in Figure 3.10.

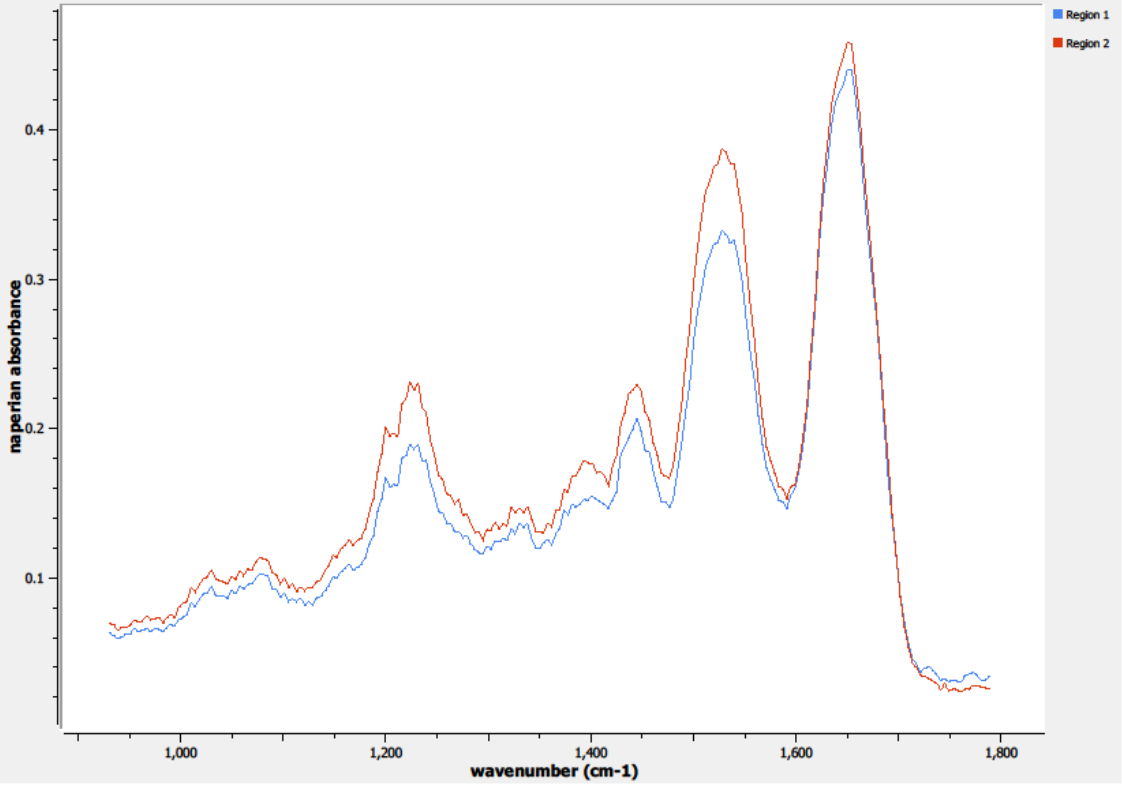



**Figure 3.10. Specimen 11-08-A085c, Point K Spectra**

#### 3.3.1.1.3. Sample 11-08-A085c, Point L

A 650 x 650 µm field-of-view image was taken at Point L on Sample 11-08-A085c using Spero's high-magnification objective (see Figure 3.11). Point L is located on a region of fat cells on the specimen, which tends to have lesser spectral information. The fat cells have a tendency to wash out during the tissue sample preparation process, which removes most of the little original spectral data from that region.

Region 1 is a general region of interest over the center of the field-of-view. Regions 2 and 3 are focusing on the cellular membranes surrounding the fat cell regions.

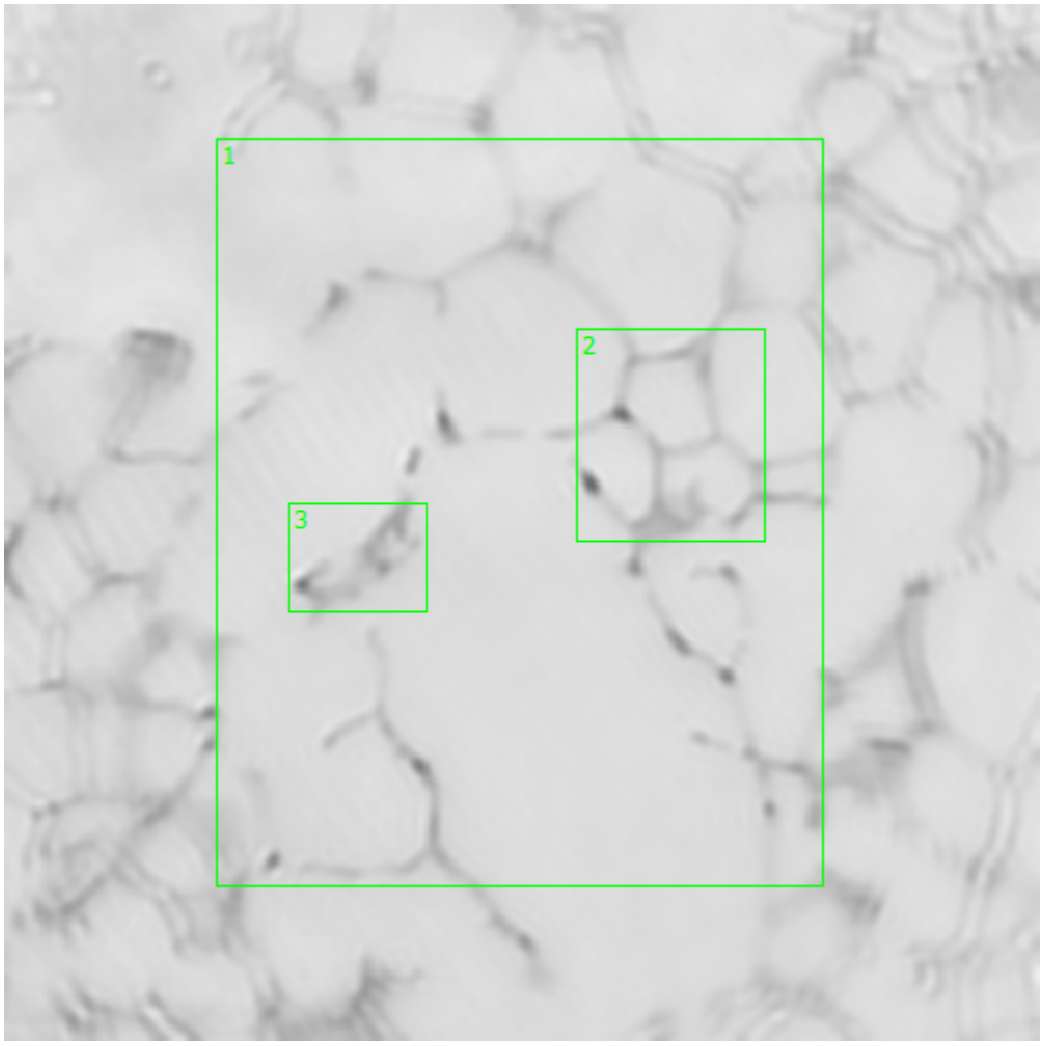



**Figure 3.11. Specimen 11-08-A085c, Point L Regions of Interest**
**Region 1: normal right breast tissue, washed out fat cells**
**Region 2, 3: concentrated regions, normal right breast tissue, washed out fat cells**

The spectra from the three defined regions of interest on Point L are depicted in Figure 3.12.

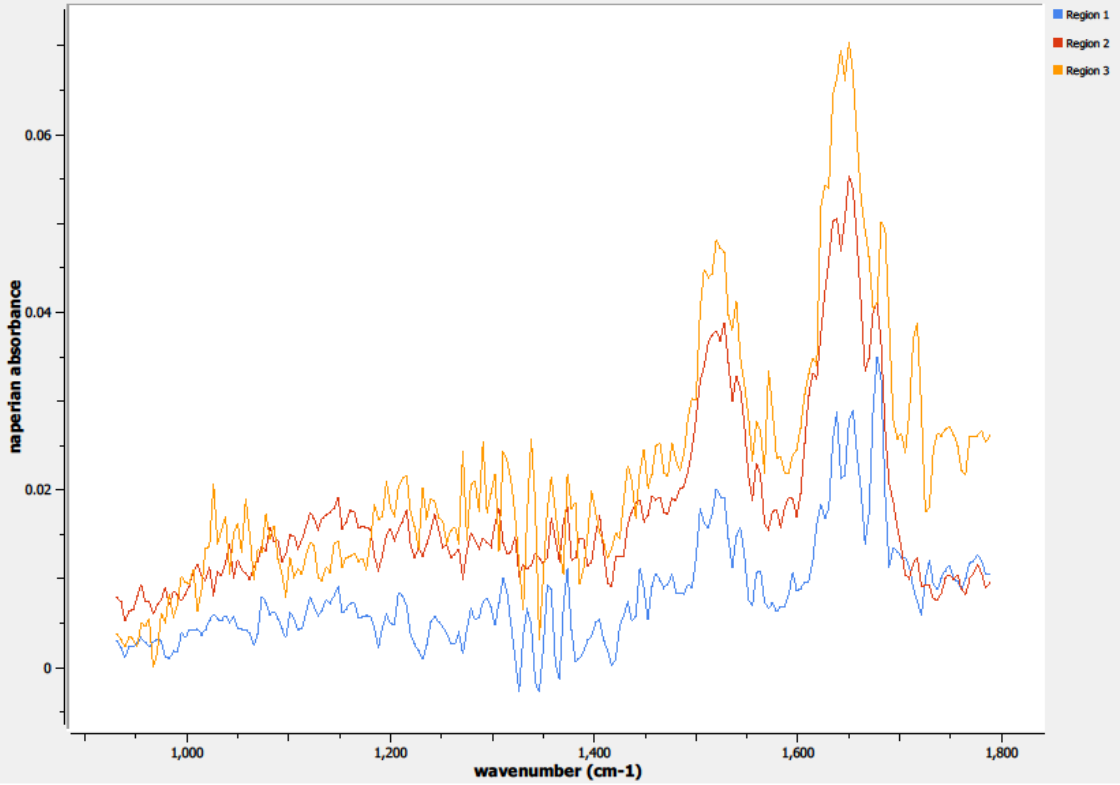



**Figure 3.12. Specimen 11-08-A085c, Point L Spectra**

### 3.3.1.2. Specimen 12-12-A095a

Specimen 12-12-A095a is a 5.80 gram sample from a 28-year-old white female. Figure 3.13 is an image of a sample of the specimen stained with H&E on a glass slide. The pink and purple areas indicate denser areas of the sample; the fat cell regions can be faintly seen in areas on the sample by the cellular membranes left from the fat cells washing out during the preparation process.

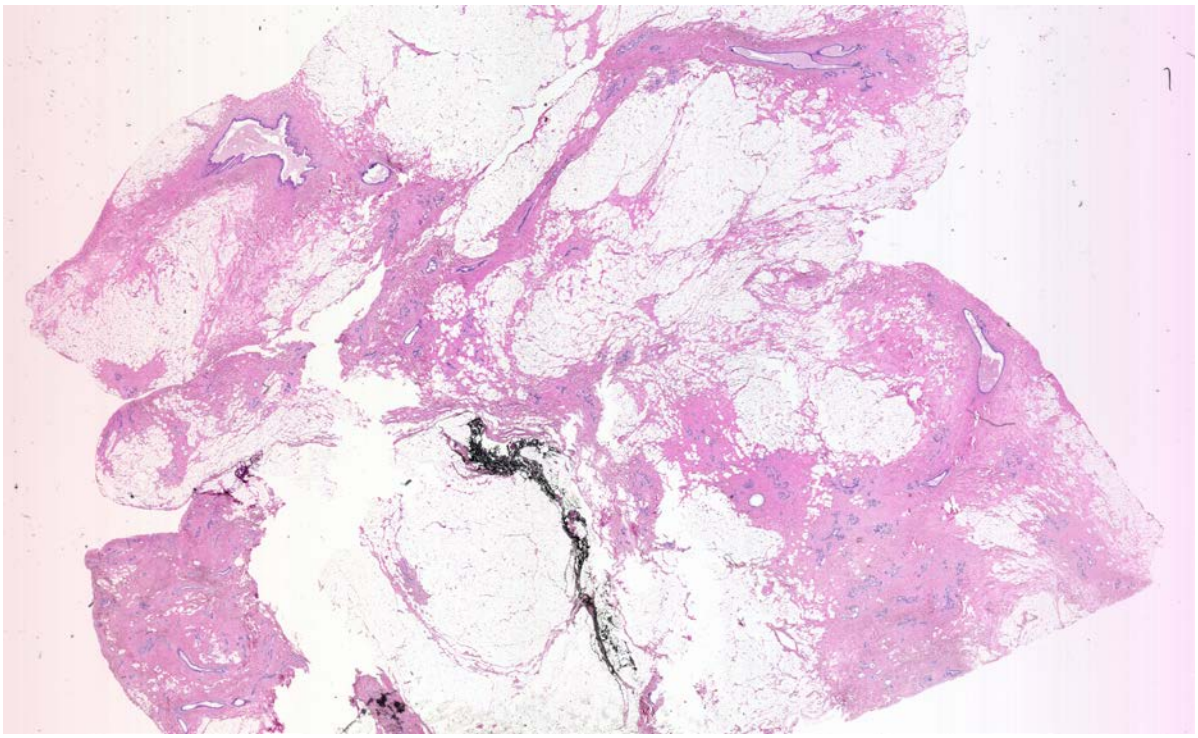


**Figure 3.13. Specimen 12-12-A095a, H&E stain, normal breast tissue.**

Figure 3.14 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flag O depicts the data collection points on Sample 12-12-A095a, focusing on a denser tissue area that will have a higher absorption.

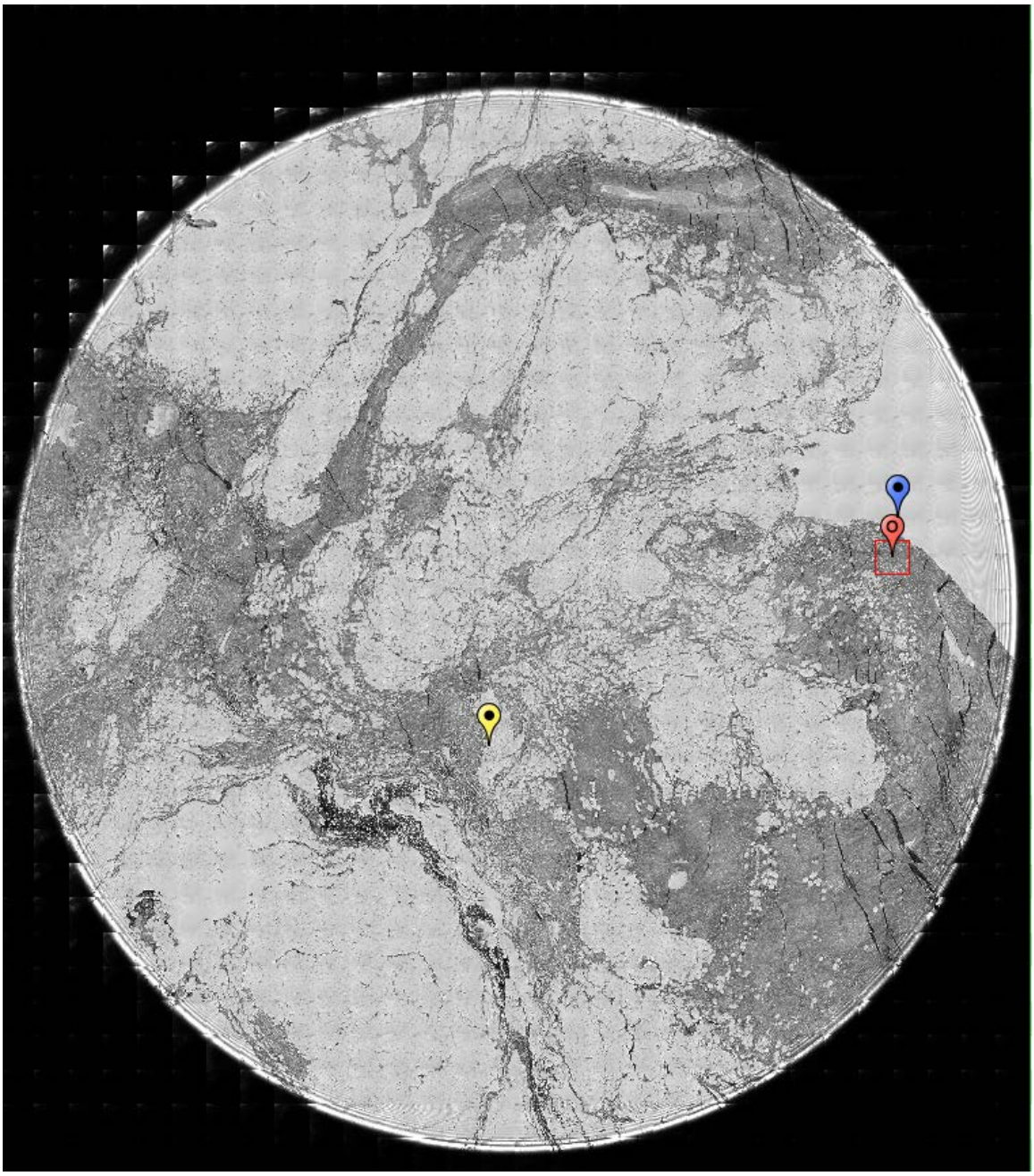


**Figure 3.14. Specimen 12-12-A095a, normal right breast tissue.**

#### 3.3.1.2.1. Sample 12-12-A095a, Point O

A 650 x 650 µm field-of-view image taken of Sample 12-12-A095a at Point O (see Figure 3.15) was obtained from Spero's high-magnification objective. Only one region of interest was needed to capture the spectral information at this point on the sample.

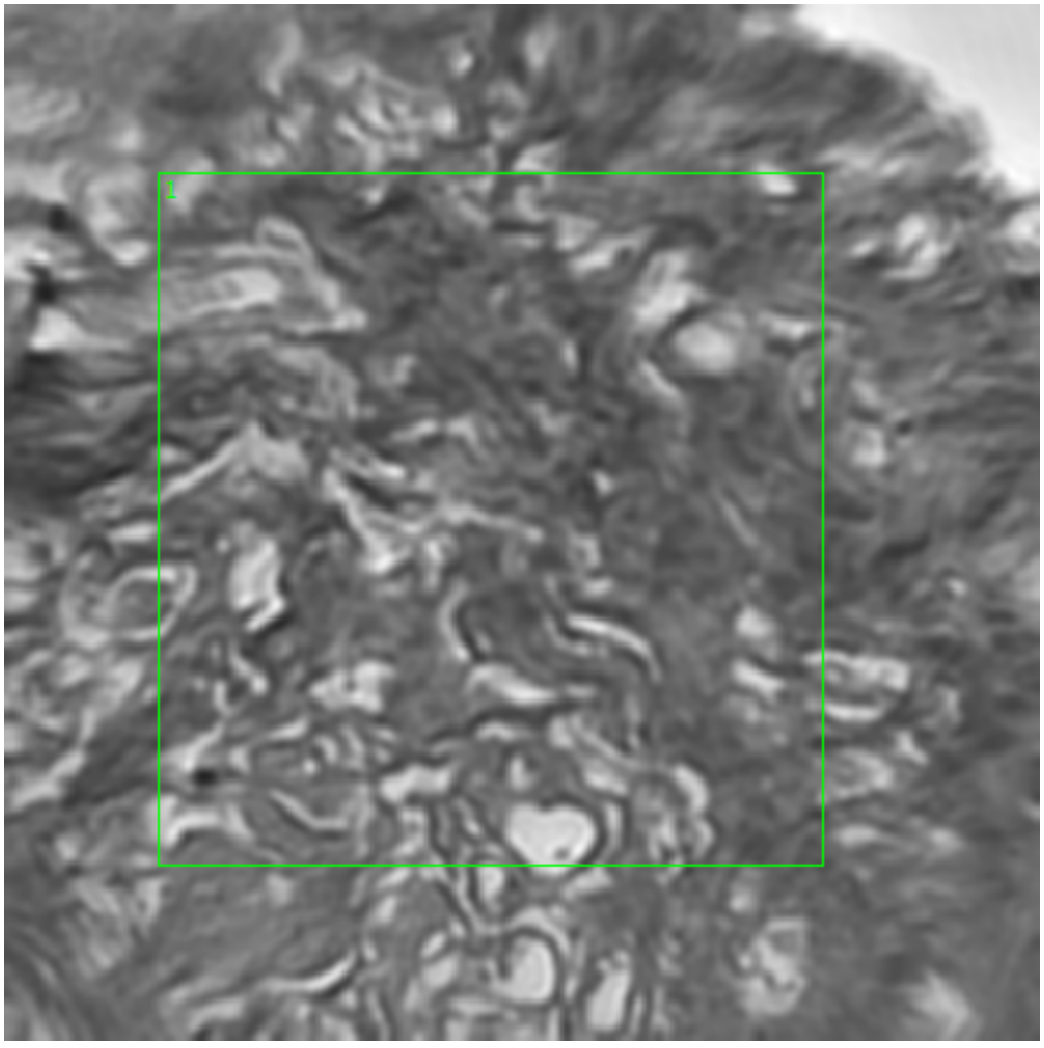



**Figure 3.15. Specimen 12-12-A095a, Point O Regions of Interest**
**Region 1: normal right breast tissue**

The spectrum from the defined region of interest on Point O is depicted in Figure 3.16.

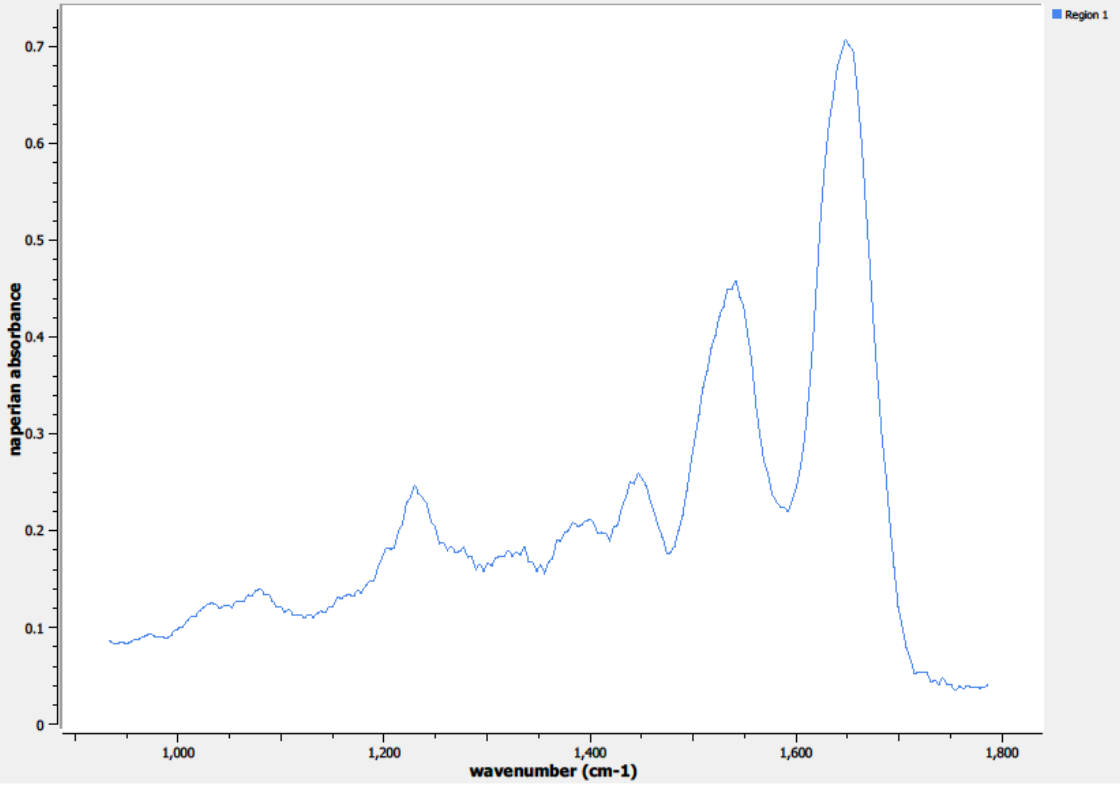



**Figure 3.16. Specimen 12-12-A095a, Point O Spectra**

### 3.3.1.3. Specimen 13-04-A173a

Specimen 13-04-A173a is a 0.37 gram sample from a 63-year-old black female. Figure 3.17 is an image of a sample of the specimen stained with H&E on a glass slide. This sample has many fat cell regions; the denser parts indicated by the pink and purple areas are not as prevalent as the previous breast examples.

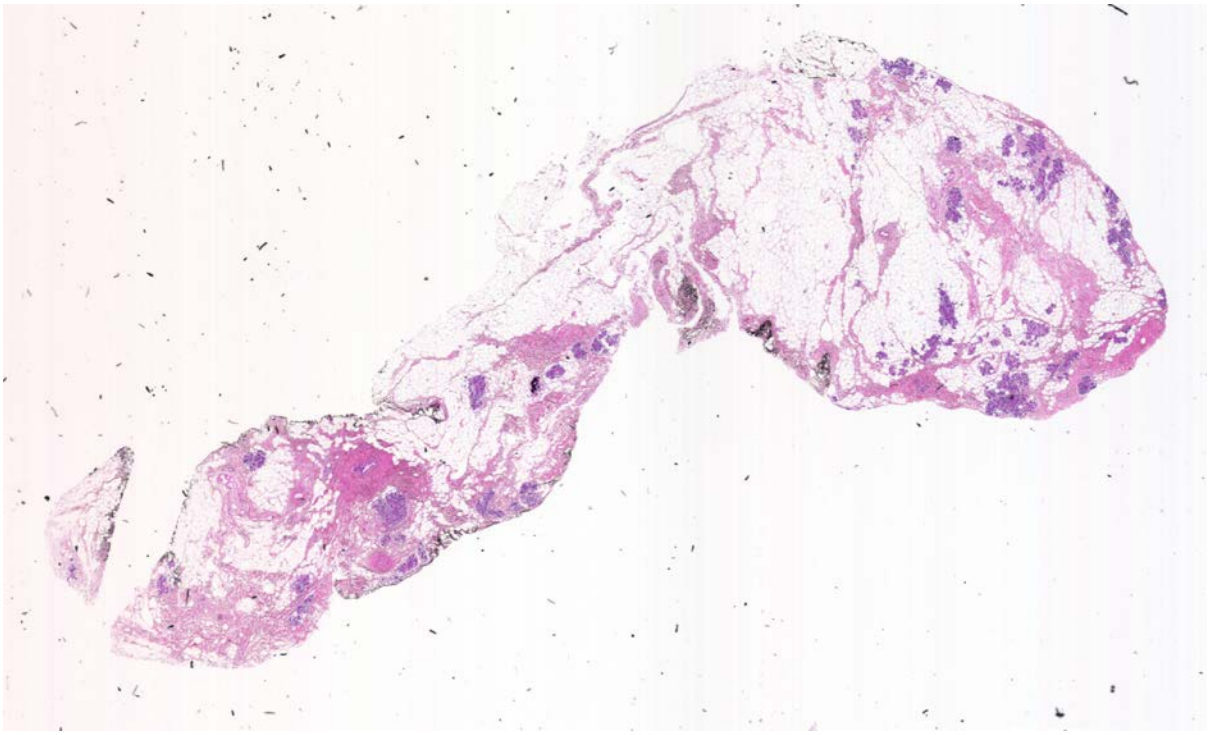


**Figure 3.17. Specimen 13-04-A173a, H&E stain, normal breast tissue.**

Figure 3.18 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags X, Y, and Z depict the data collection points on Sample 13-04-A173a, focusing on the denser tissue areas that will have a higher absorption.

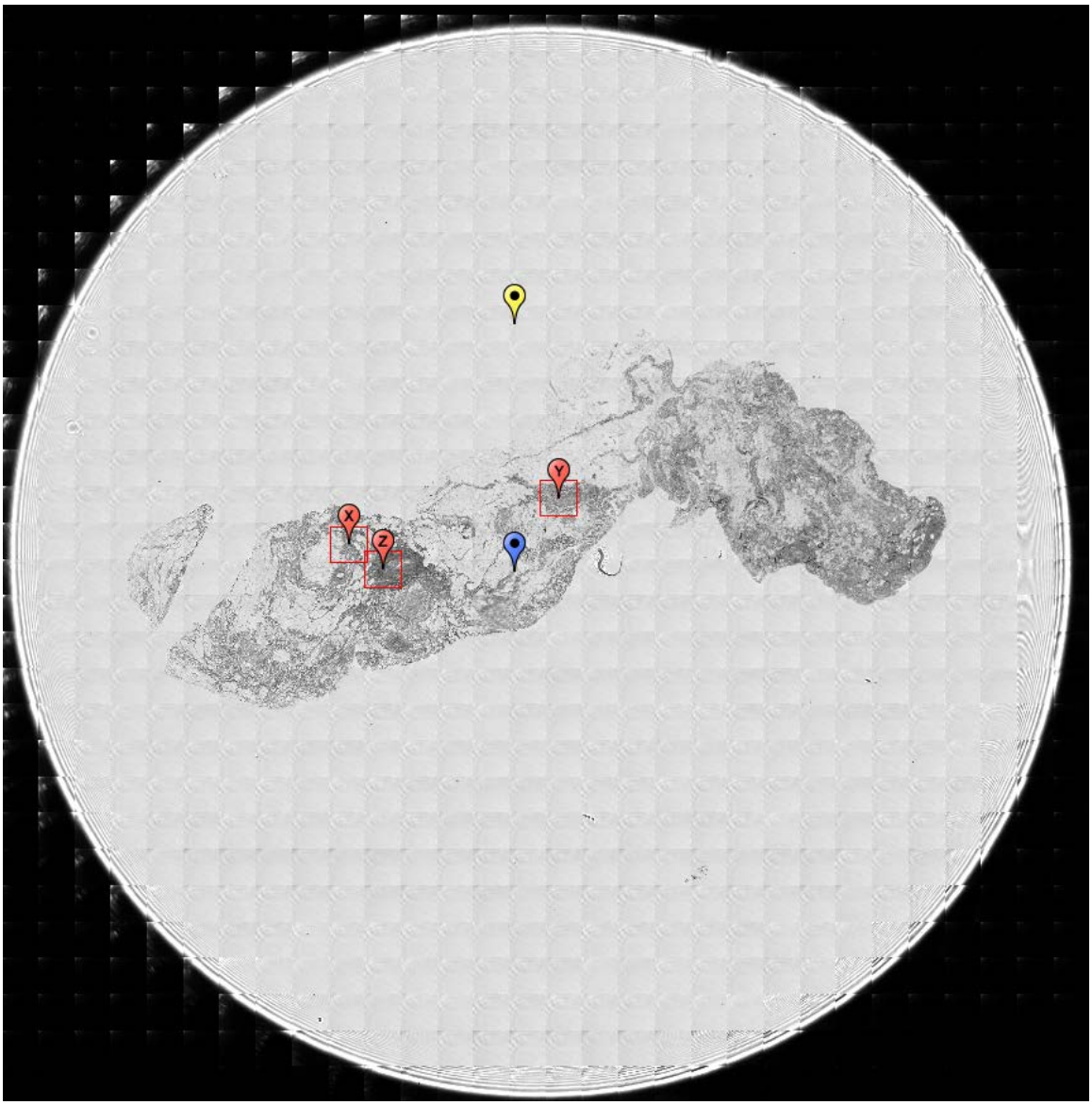



**Figure 3.18. Specimen 13-04-A173a, normal breast tissue.**

#### 3.3.1.3.1. Sample 13-04-A173a, Point X

A 650 x 650 μm field-of-view image taken of Sample 13-04-A173a at Point X (see Figure 3.19) was obtained from Spero's high-magnification objective. Regions

1 and 2 are adjacent to washed-out fat cell sections and are defining more dense tissue areas for spectral data collection.

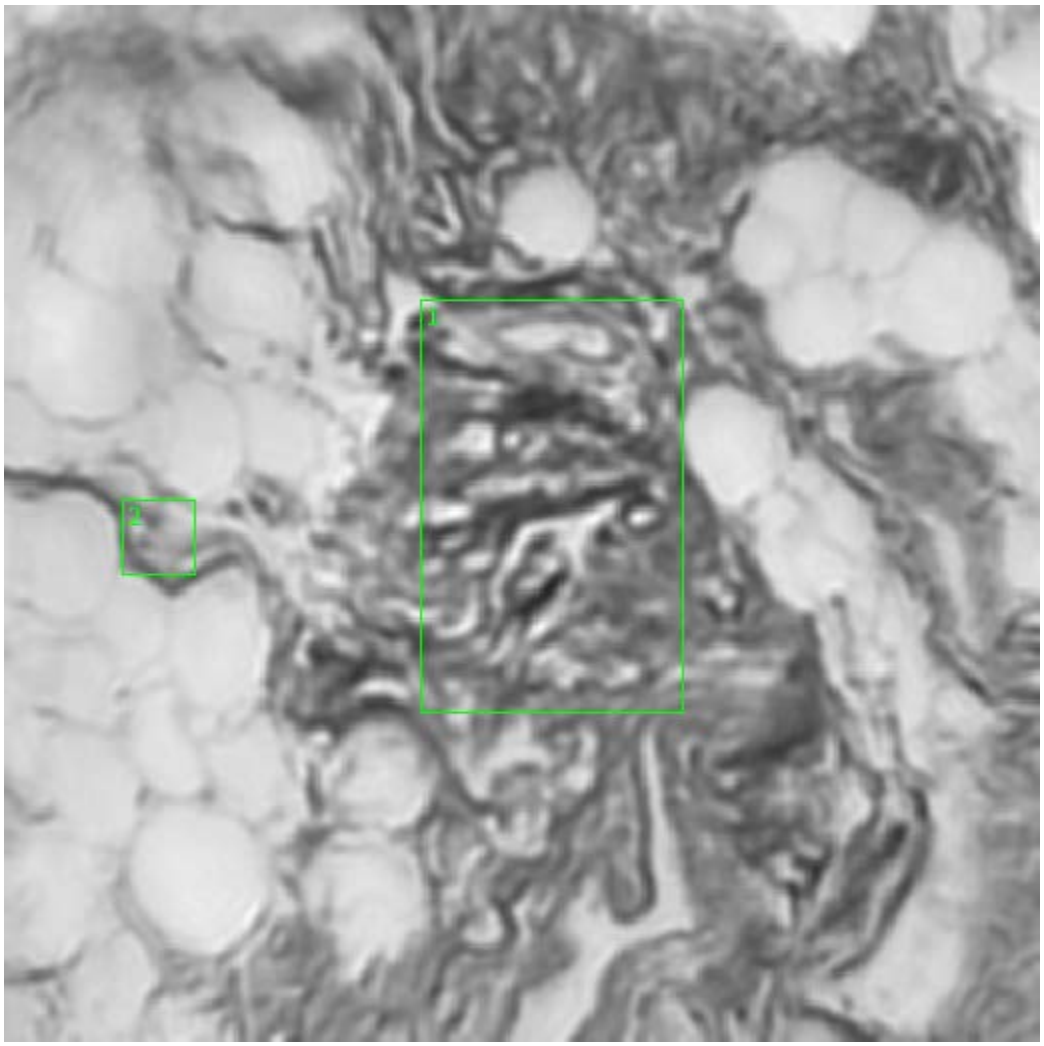



**Figure 3.19. Specimen 13-04-A173a, Point X Regions of Interest**
**Regions 1, 2: normal breast tissue**

The spectra from the two defined regions of interest on Point X are portrayed in Figure 3.20.

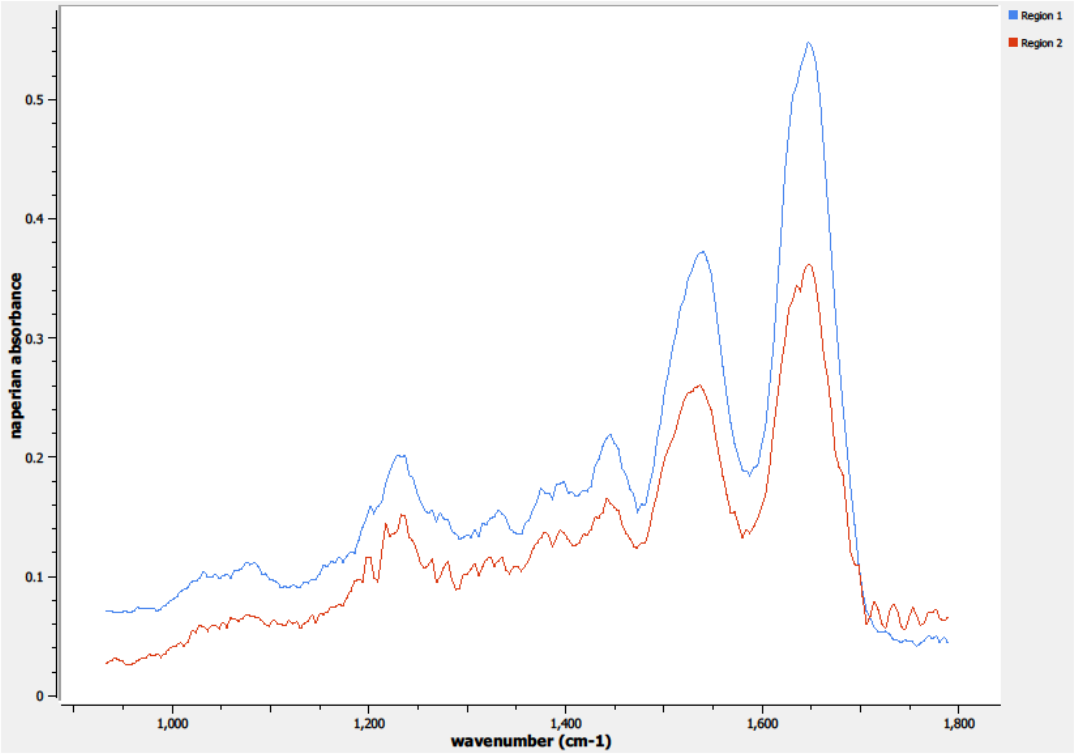



**Figure 3.20. Specimen 13-04-A173a, Point X Spectra**

### 3.3.1.3.2. Sample 13-04-A173a, Point Y

The high-magnification 650 x 650 µm field-of-view image was also taken at Point Y (see Figure 3.21) on Sample 13-04-A173a. Only one region of interest was needed to capture the spectral information at this point on the sample.

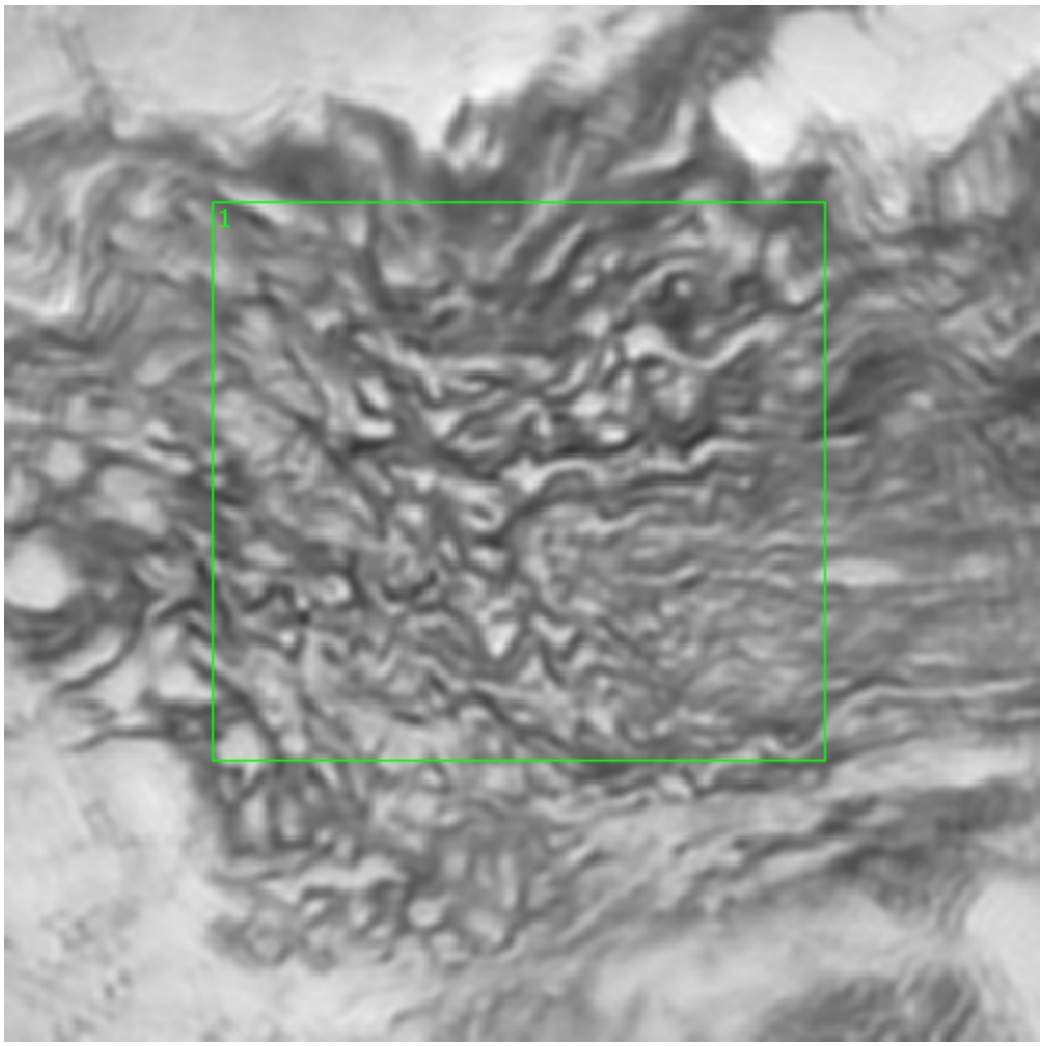



**Figure 3.21. Specimen 13-04-A173a, Point Y Regions of Interest**
**Region 1: normal breast tissue**

The spectrum from the defined region of interest on Point Y is shown in Figure 3.22.

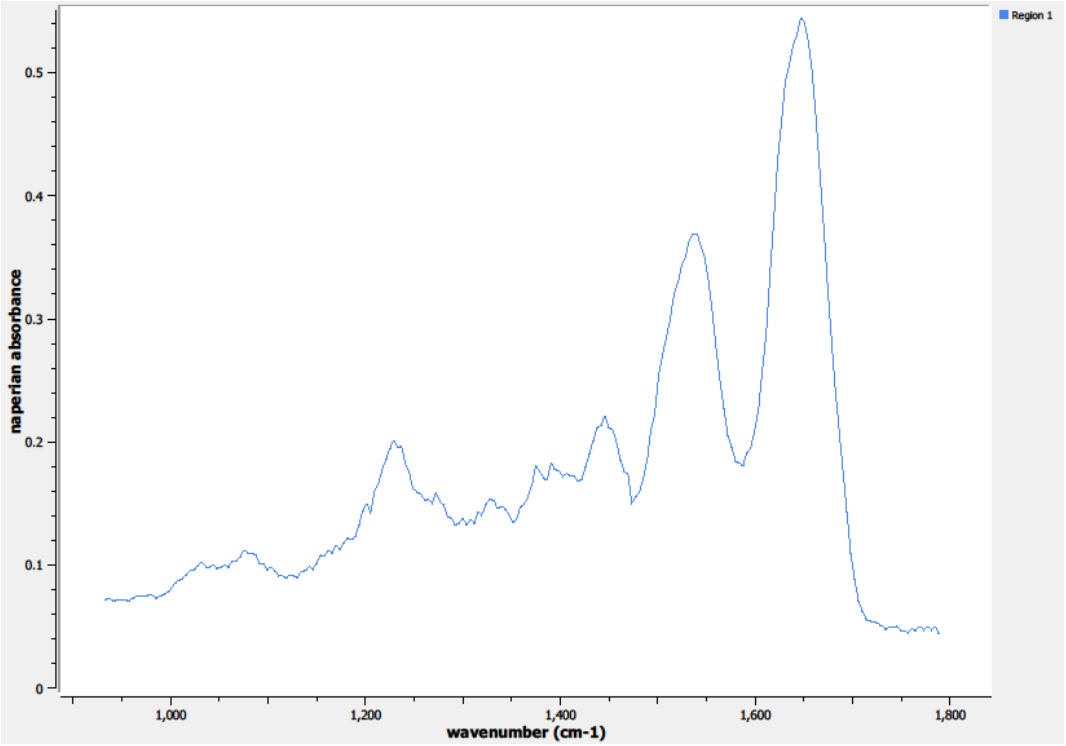



**Figure 3.22. Specimen 13-04-A173a, Point Y Spectra**

#### 3.3.1.3.3. Sample 13-04-A173a, Point Z

The high-magnification field-of-view image taken at Point Z on Sample 13-04-A173a can be seen in Figure 3.23. Regions 1 and 6 are defining more dense tissue areas for spectral data collection. Regions 2 and 3 are focusing on an area adjacent to a glandular structure. The area is the tissue surrounding and defining the gland and is embedded in the denser tissue regions. Region 5 is defining a region of interest interior to a glandular structure; low absorption is anticipated in Region 5.

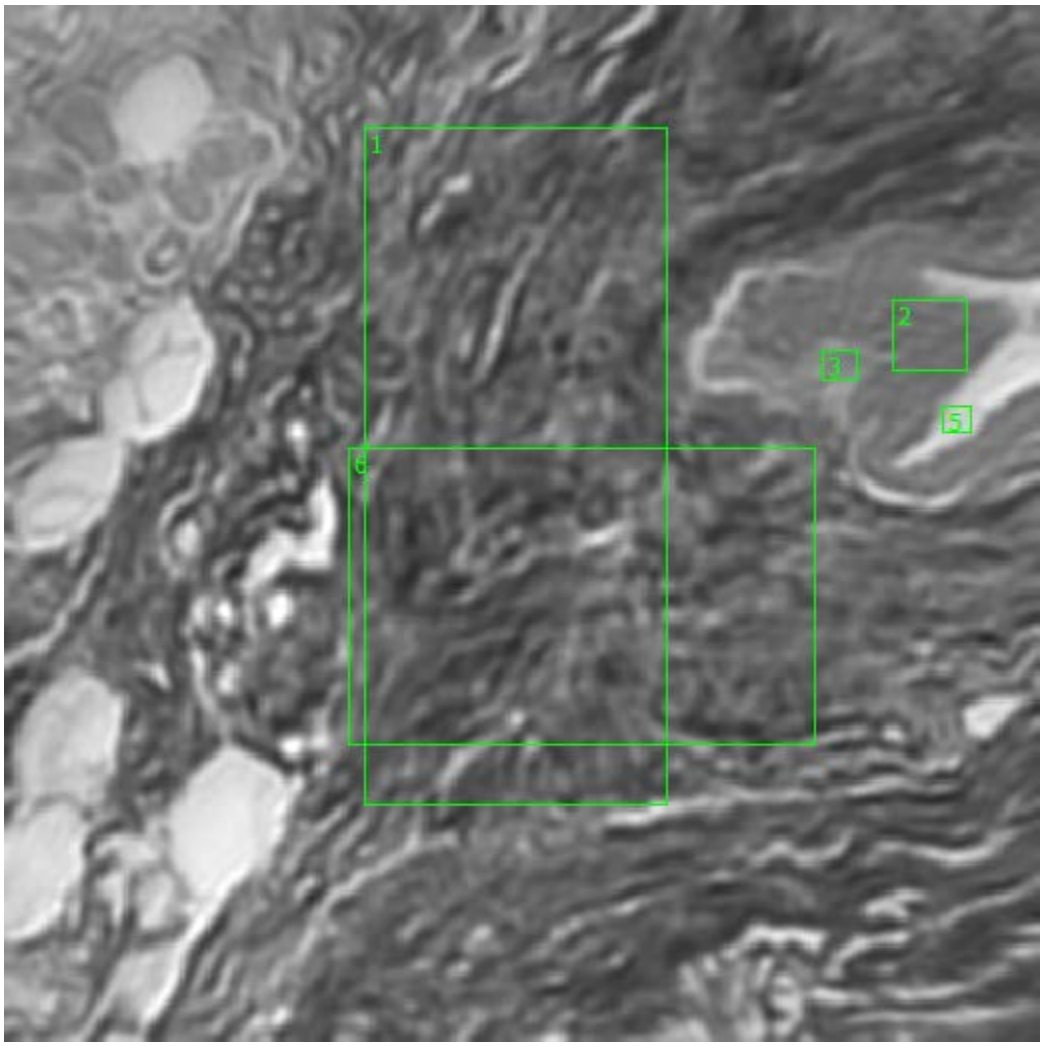



**Figure 3.23. Specimen 13-04-A173a, Point Z Regions of Interest**
**Regions 1, 6: normal breast tissue**
**Regions 2, 3: membrane surrounding glandular structure**
**Region 5: interior glandular structure**

The spectra from the five defined regions of interest on Point Z are depicted in Figure 3.24.

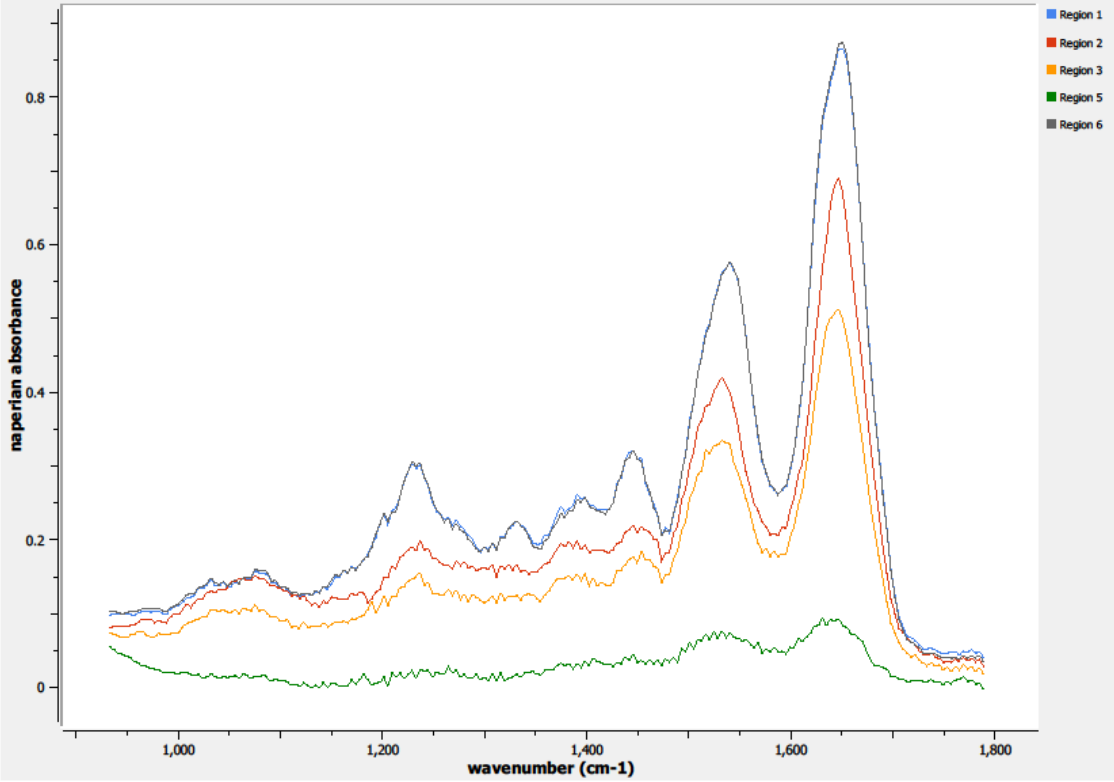



**Figure 3.24. Specimen 13-04-A173a, Point Z Spectra**

### 3.3.1.4. Specimen 13-02-A117a

Specimen 13-02-A117a is a 1.31 gram sample from a 60-year-old white female. Figure 3.25 is an image of a sample of the specimen stained with H&E on a glass slide.

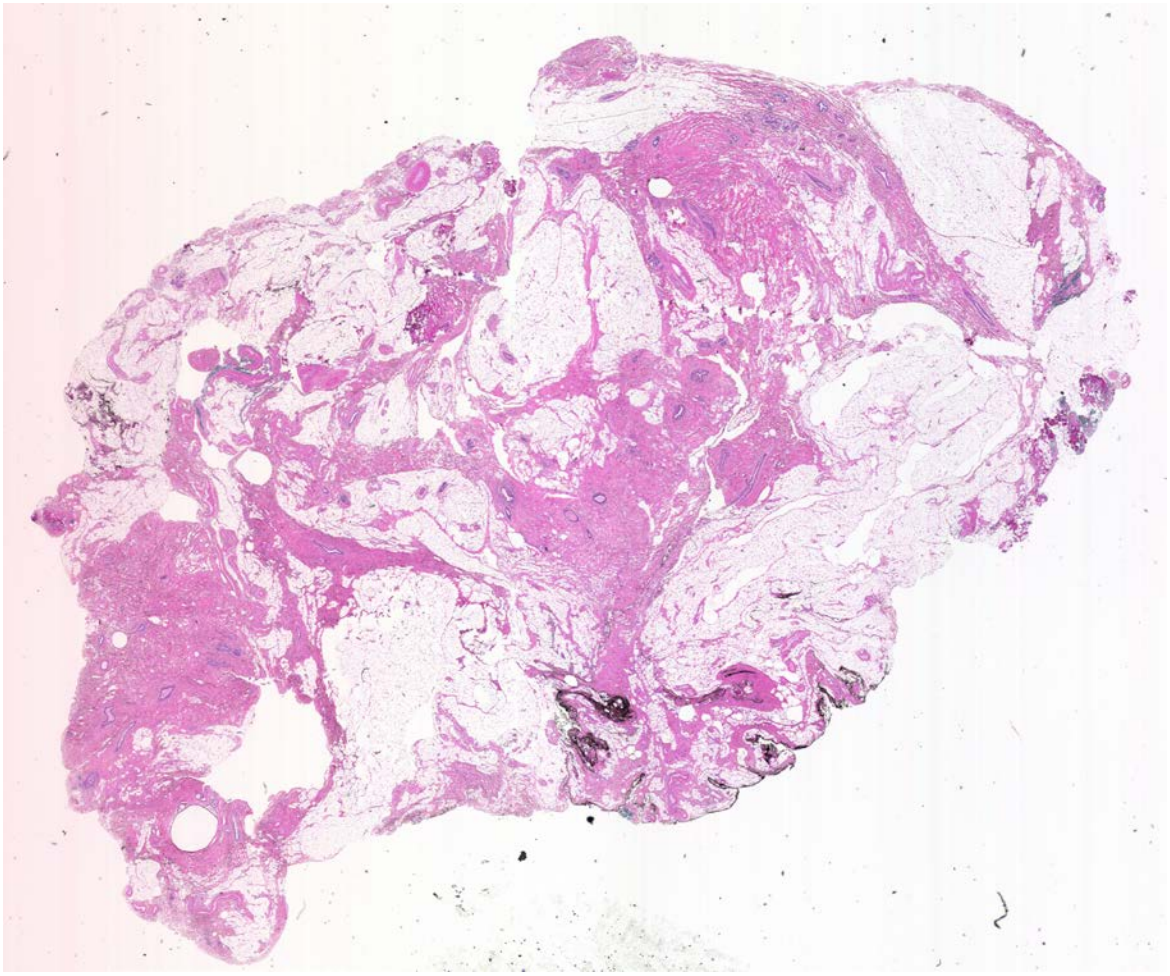


**Figure 3.25. Specimen 13-02-A117a, H&E stain, normal breast tissue.**

Figure 3.26 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags T and U depict the data collection points on Sample 13-02-A117a, focusing on the denser tissue areas that will have a higher absorption.

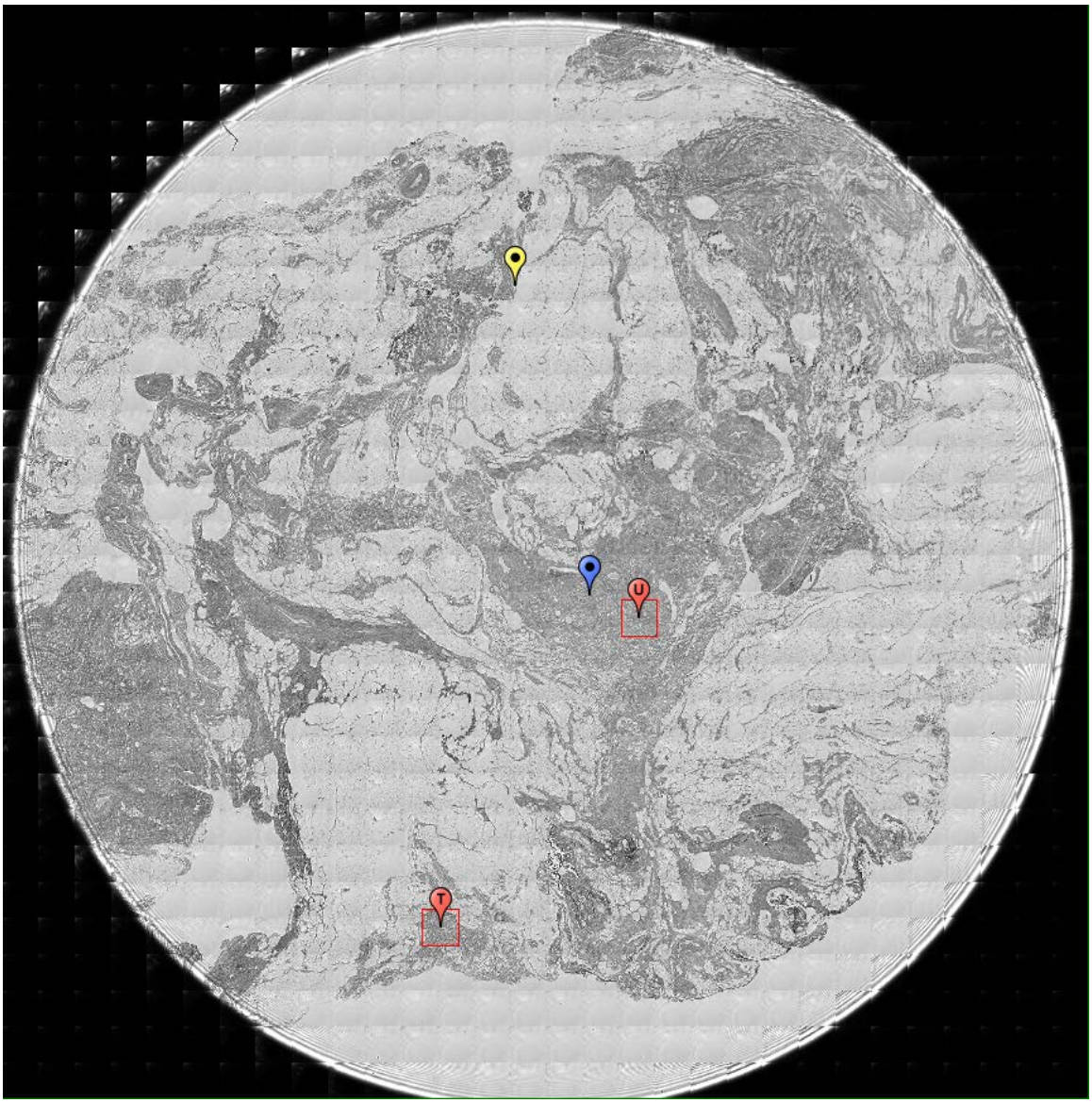



**Figure 3.26. Specimen 13-02-A117a, normal breast tissue.**

#### 3.3.1.4.1. Sample 13-02-A117a, Point T

A 650 x 650 μm field-of-view image taken of Sample 13-02-A117a at Point T (see Figure 3.27) was obtained from Spero's high-magnification objective. Only one

region of interest was needed to capture the spectral information at this point on the sample.

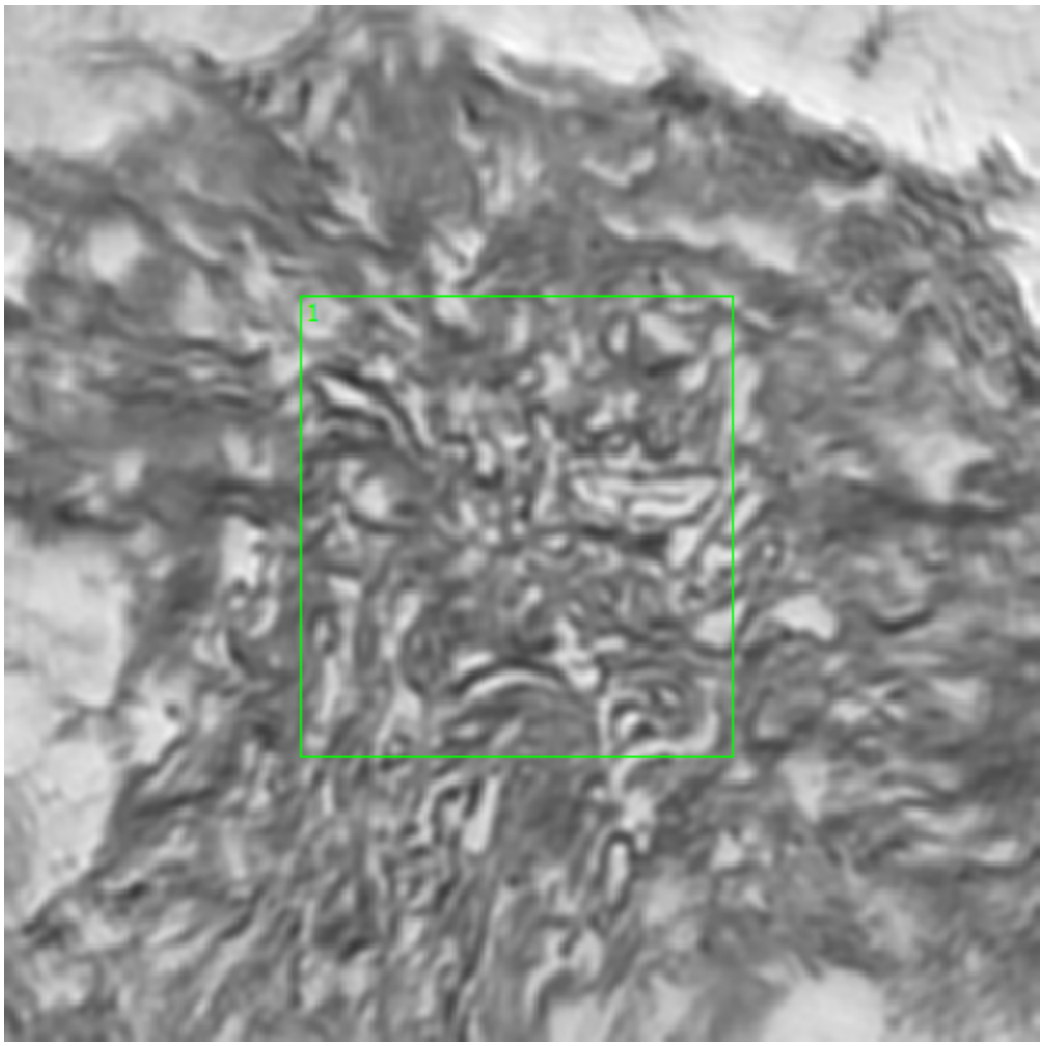



**Figure 3.27. Specimen 13-02-A117a, Point T Regions of Interest**
**Region 1: normal breast tissue**

The spectrum from the defined region of interest on Point T is portrayed in Figure 3.28.

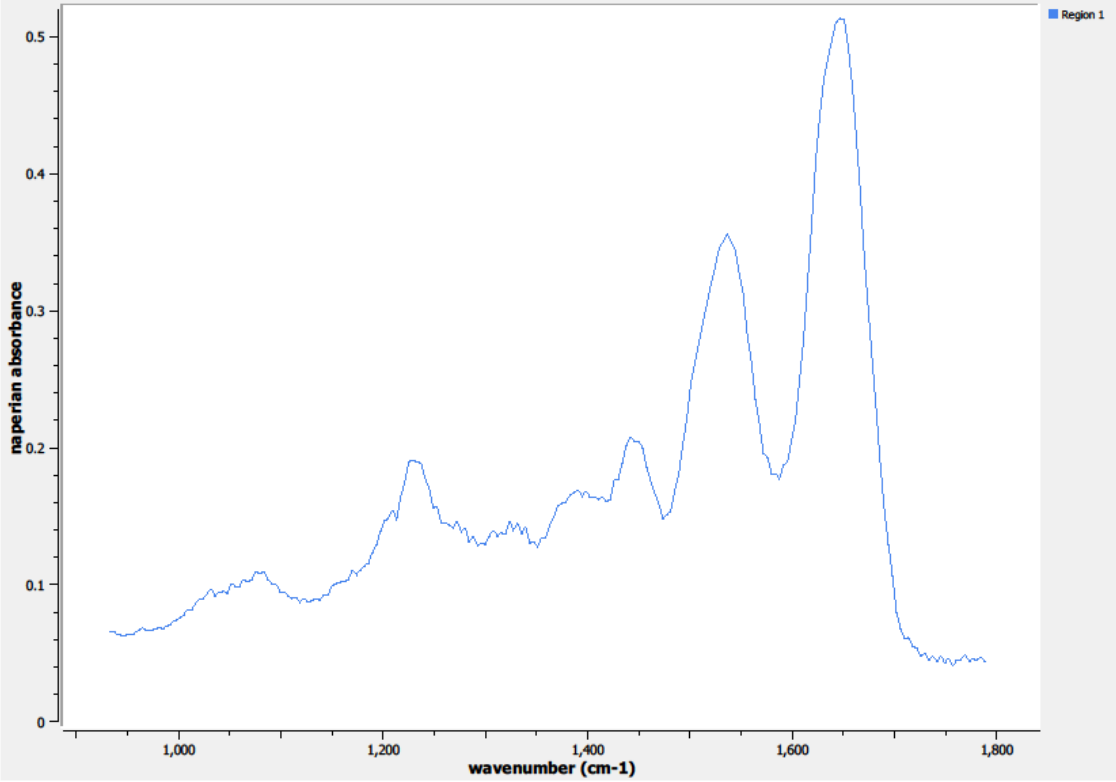



**Figure 3.28. Specimen 13-02-A117a, Point T Spectra**

#### 3.3.1.4.2. Sample 13-02-A117a, Point U

The high-magnification field-of-view image taken at Point U on Sample 13-02-A117a can be seen in Figure 3.29. Only one region of interest was needed to capture the spectral information at this point on the sample.

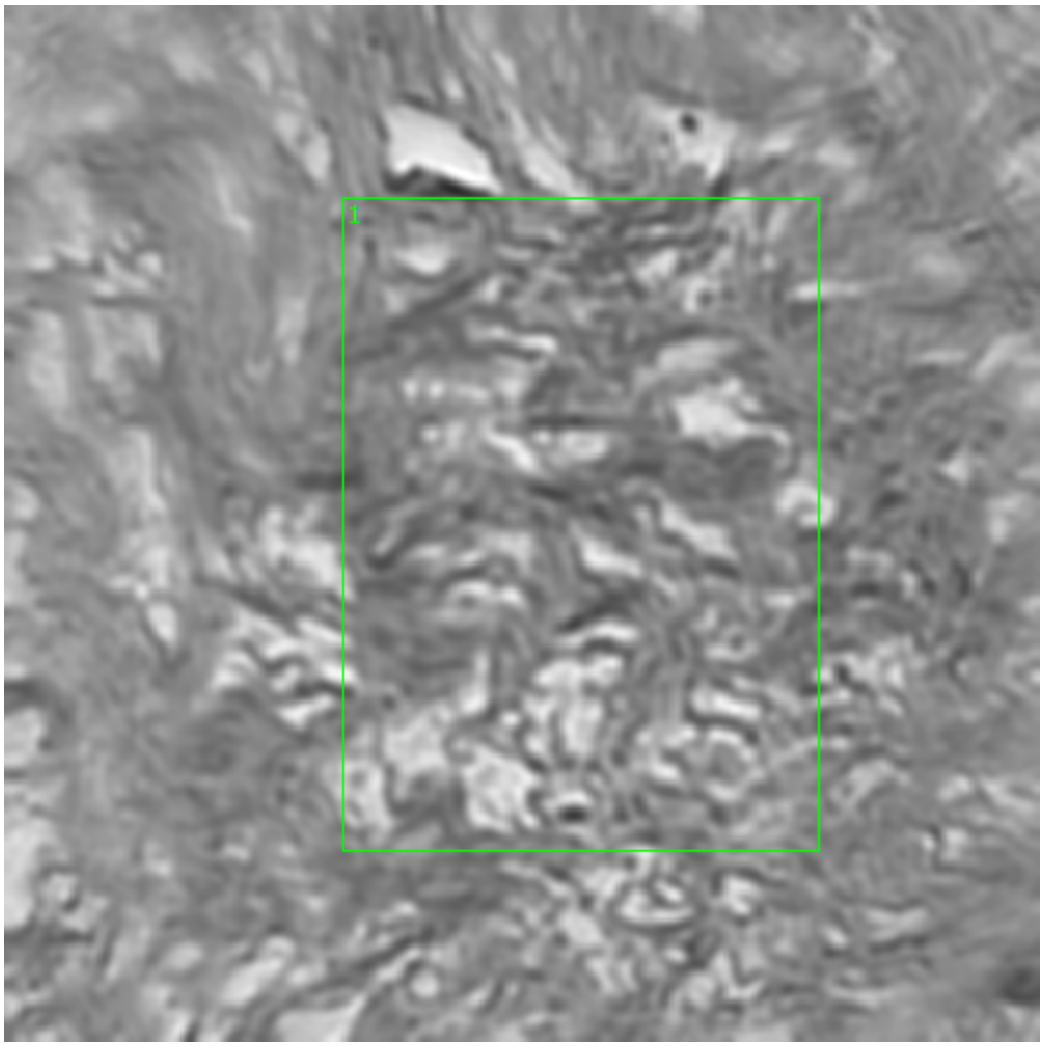



**Figure 3.29. Specimen 13-02-A117a, Point U Regions of Interest**
**Region 1: normal breast tissue**

The spectrum from the defined region of interest on Point U is depicted in Figure 3.30.

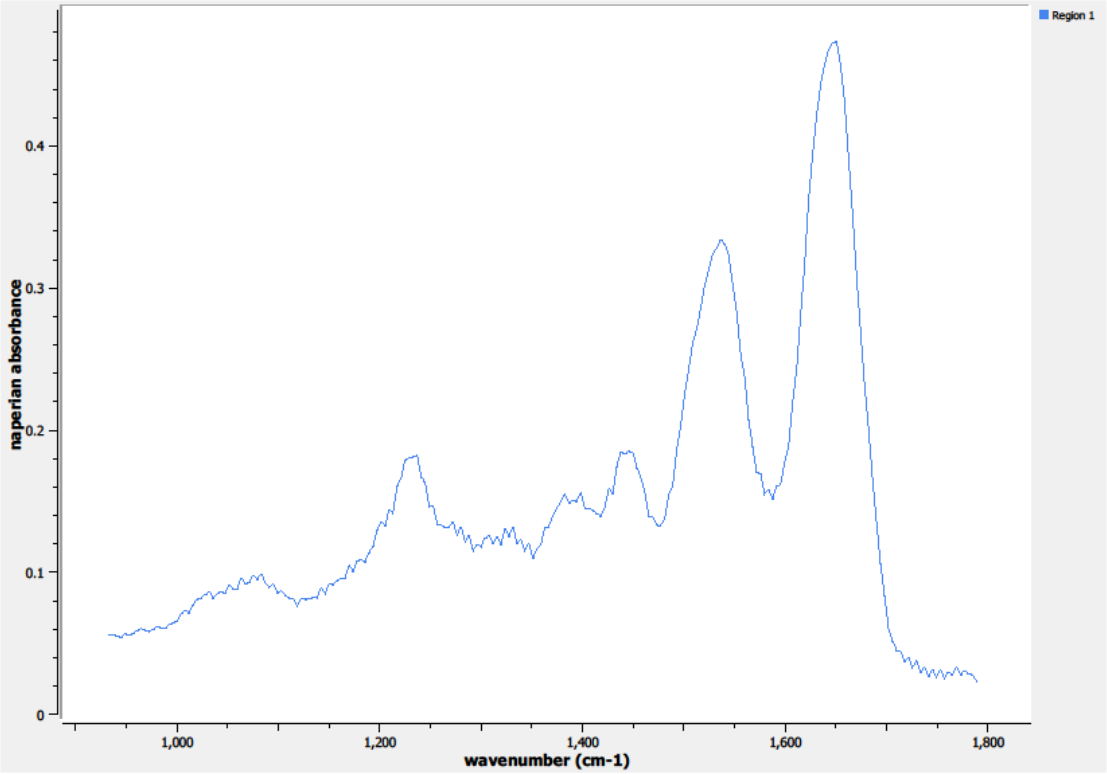



**Figure 3.30. Specimen 13-02-A117a, Point U Spectra**

### 3.3.2. Malignant Breast Data Acquisition

Fifteen malignant breast specimens were utilized for this experiment. Samples from each specimen were placed on glass slides and CaF2 windows; the samples on the glass slides were stained with H&E. Data for each sample on the CaF2 windows were collected on Spero and evaluated in this experiment; the locations of the data collection points were first identified on the samples stained with H&E. Each sample had multiple points of data collection; emphasis was placed

on areas of denser tissue indicative of tumor growth. A few key malignant breast specimen examples are listed in this section. Refer to Appendix D for a complete listing of malignant breast specimens, including images from the H&E-stained slides, the images obtained through Spero's high-magnification objective, and the spectra from each data point collected.

#### 3.3.2.1. Specimen 13-05-A013h

Specimen 13-05-A013h is a malignant breast sample from a 64-year-old white female. The specimen is a 0.98 gram primary malignant breast sample taken from the ductal region. The quality control diagnosis provided by CHTN personnel classifies the specimen as invasive malignant ductal adenocarcinoma of the breast. The tumor percentage for Specimen 13-05-A013h is identified as 100% tumor.

The hormonal receptors are diagnosed by CHTN on this specimen as estrogen positive and progesterone negative. The protein receptor HER2 is identified as negative.

An image of a sample from Specimen 13-05-A013h stained with H&E on a glass slide is shown in Figure 3.31. The dark purple throughout the image is indicative of high cell concentration typically associated with a tumor growth. The invasiveness of the tumor is evident throughout the entire tissue sample.

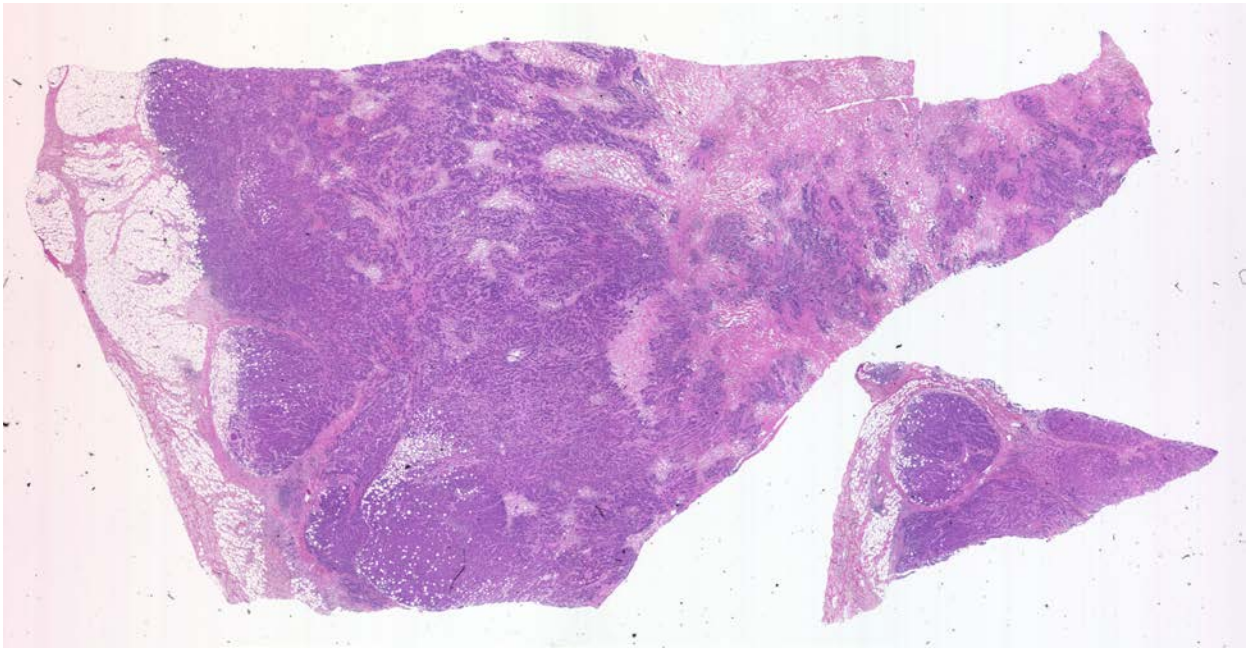


**Figure 3.31. Specimen 13-05-A013h, H&E stain, malignant ductal adenocarcinoma of the breast**

Figure 3.32 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags M, N, P, Q, and S depict the spectral data collection points on Sample 13-05-A013h, focusing on areas of higher cell concentration. Because the flags are duplicating data collection for similar regions, we only look at a flags N, P, and S in this section. Refer to Appendix D for the images and spectral data of the remaining malignant regions specified by the flag locations seen in Figure 3.32.

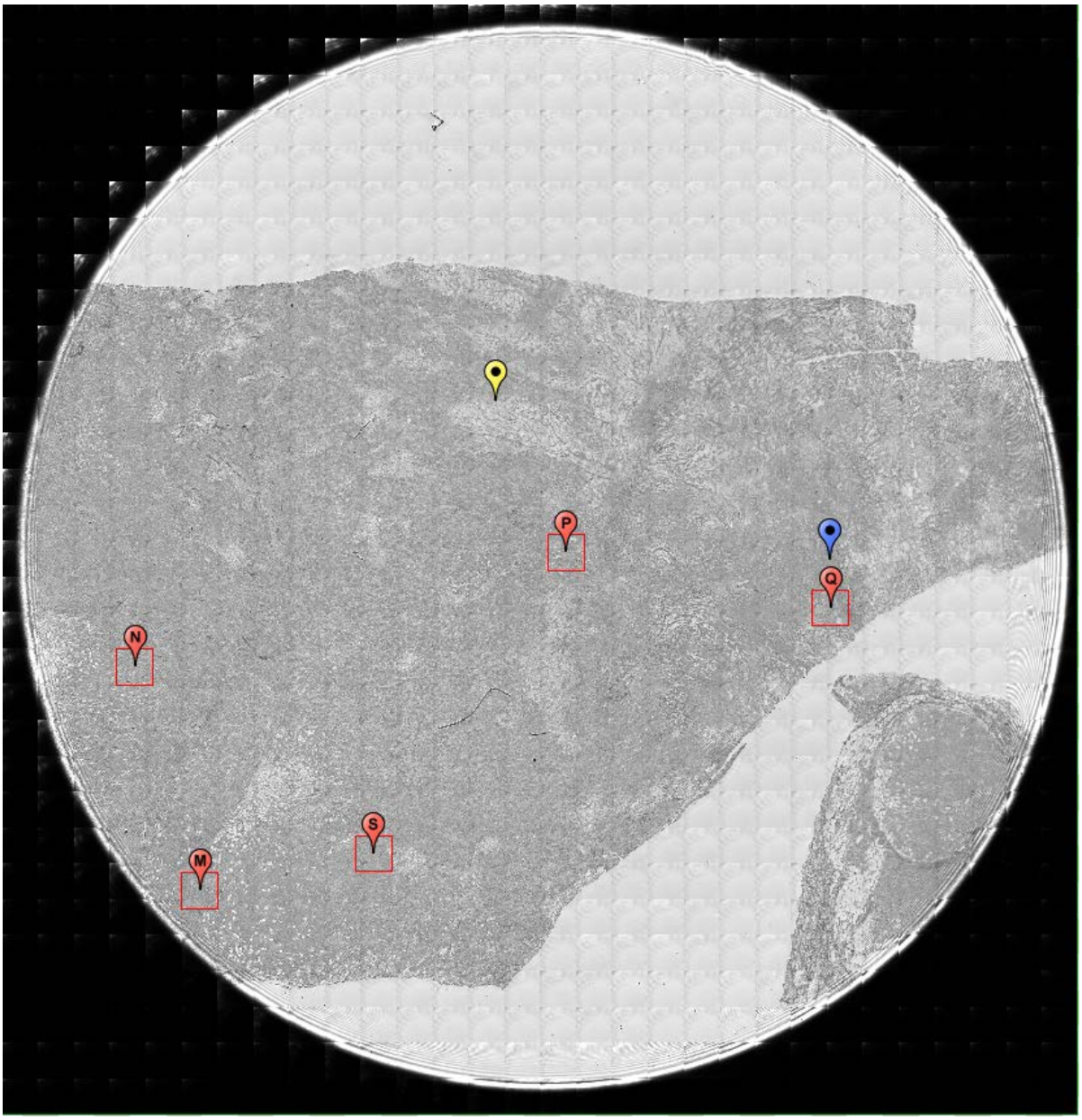



**Figure 3.32. Specimen 13-05-A013h, ductal breast cancer tissue**

#### 3.3.2.1.1. Sample 13-05-A013h, Point N

A 650 x 650 μm field-of-view image taken of the specimen at Point N (see Figure 3.33) was obtained from Spero's high-magnification objective. Point N's location on Sample 13-05-A013h is targeting one of the sample's higher cell concentration areas, depicted by the dark purple on the H&E stained image. Two regions of interest were collected due to the missing tissue in the region which is typical of breast samples. Region 1 is collecting spectral information over the majority of the field-of-view, while Region 2 is focused on a smaller area with more concentrated tissue.

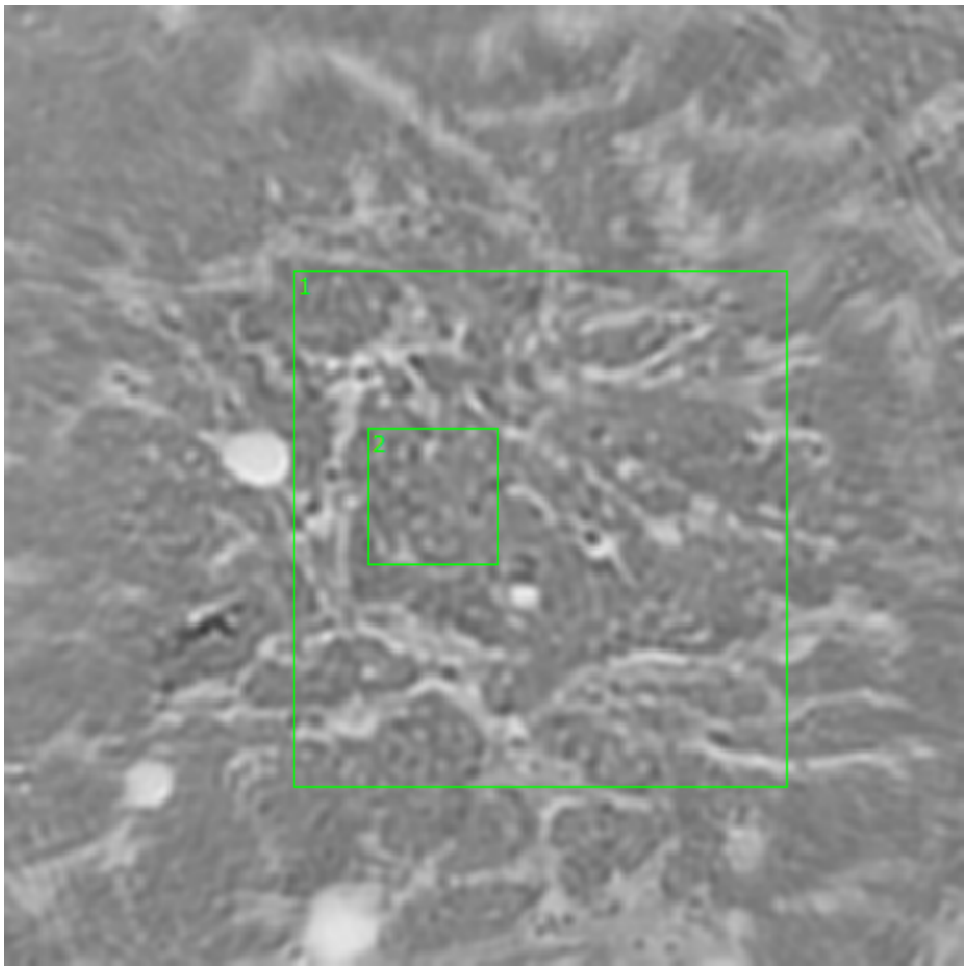



**Figure 3.33. Specimen 13-05-A013h, Point N Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**
**Region 2: breast, malignant ductal adenocarcinoma, concentrated region**

The spectrum from the defined region of interest on Point N is shown in Figure 3.34. As is seen throughout the breast specimens, the mid-infrared absorbance is often much lower with the breast tissue than the denser and more muscular cervical tissue studied in the previous chapter.

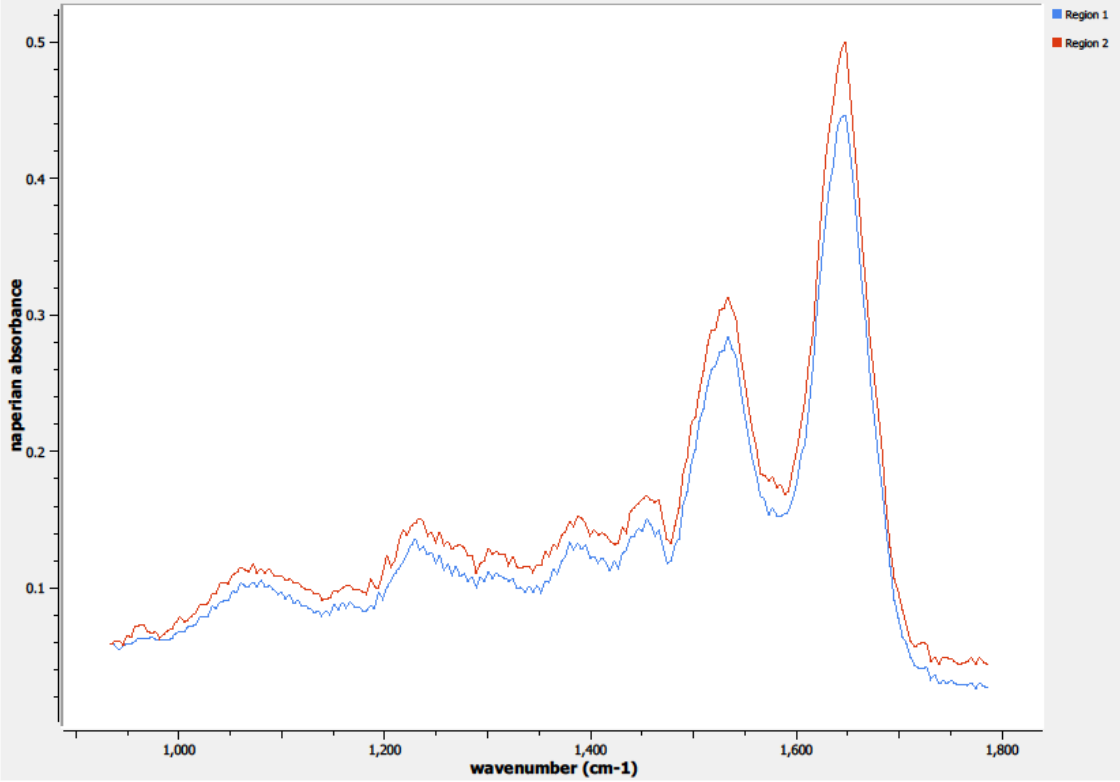



**Figure 3.34. Specimen 13-05-A013h, Point N Spectra**

#### 3.3.2.1.2. Sample 13-05-A013h, Point P

The high-magnification field-of-view image taken at Point P on Sample 13-05-A013h can be seen in Figure 3.35. Point P is located on the interior of the sample, again targeting one of the sample's higher cell concentration areas. Similar to Point N, two regions of interest were collected due to the missing tissue in the region which is typical of breast samples. Region 1 is collecting spectral information over the majority of the field-of-view, while Region 2 is focused on a smaller area with more concentrated tissue.

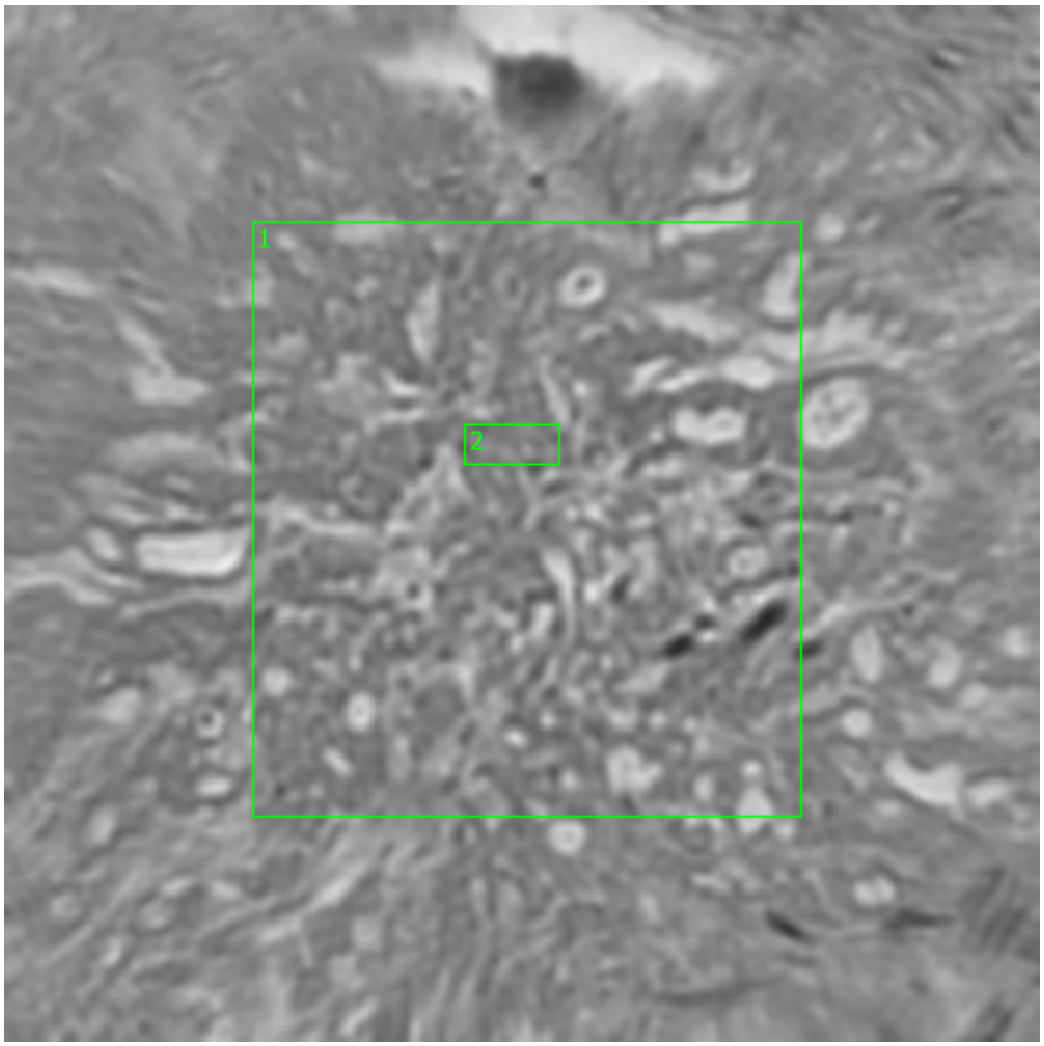



**Figure 3.35. Specimen 13-05-A013h, Point P Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**
**Region 2: breast, malignant ductal adenocarcinoma, concentrated region**

The spectrum from the defined region of interest on Point P is depicted in Figure 3.36. As anticipated due to the small region of interest, Region 2 has more noise in the spectrum than Region 1.

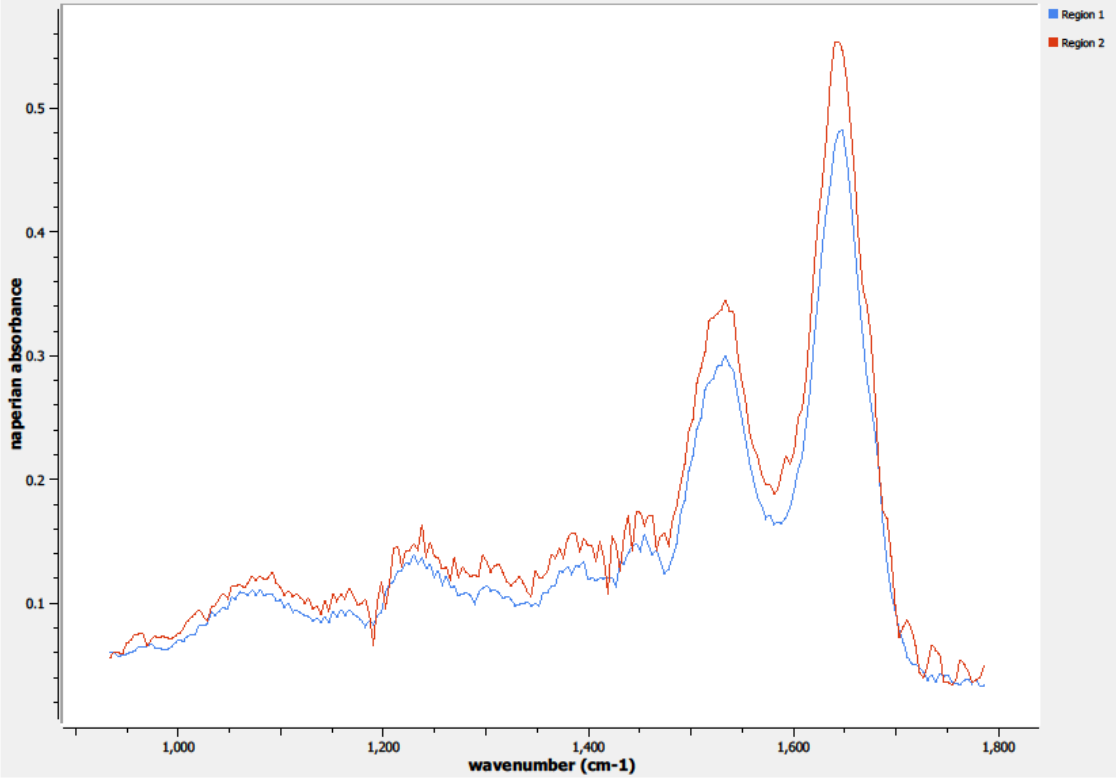



**Figure 3.36. Specimen 13-05-A013h, Point P Spectra**

#### 3.3.2.1.3. Sample 13-05-A013h, Point S

The high-magnification field-of-view image taken at Point S on Sample 13-05-A013h can be seen in Figure 3.37. Point S is again targeting one of the sample's higher cell concentration areas. Only one region of interest was needed to capture the spectral information at Point S.

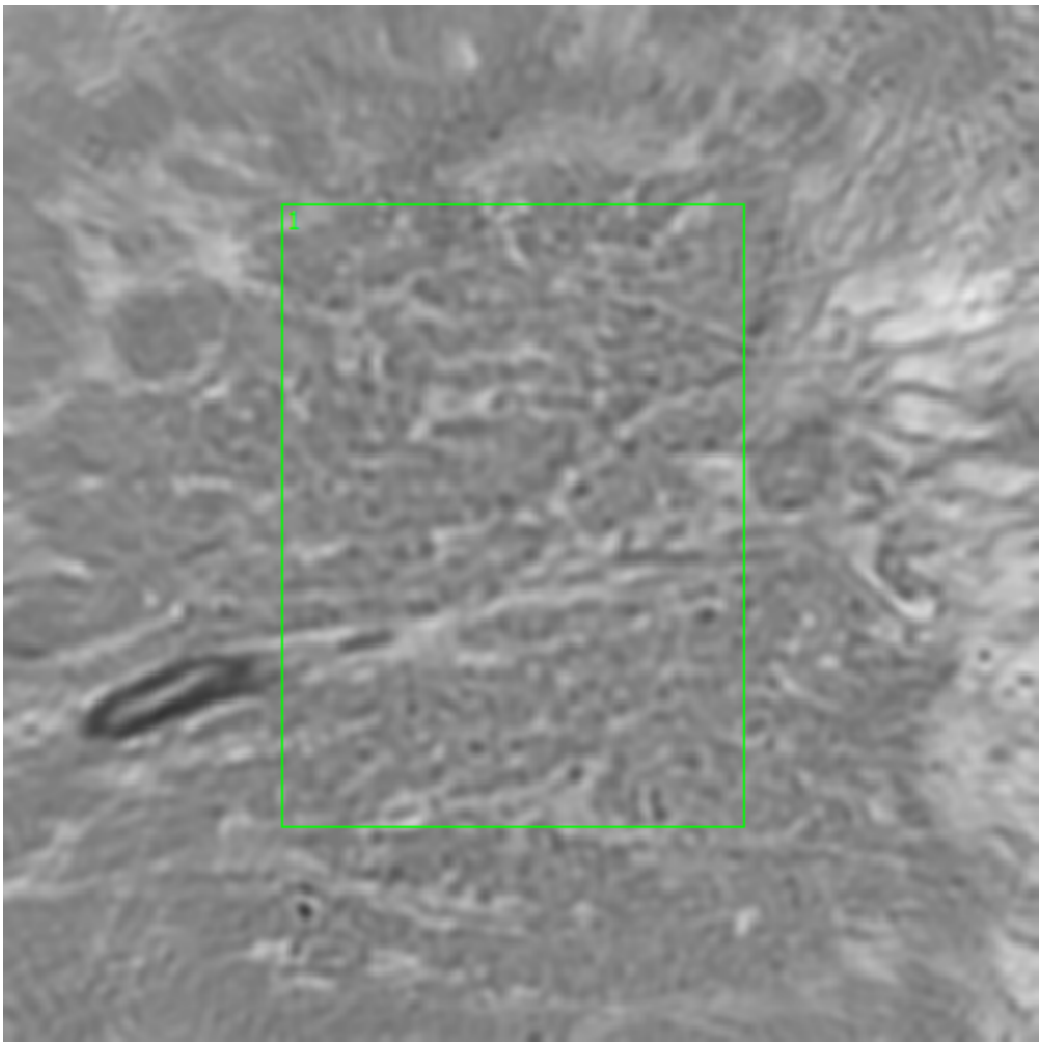



**Figure 3.37. Specimen 13-05-A013h, Point S Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

The spectrum from the defined region of interest on Point S is depicted in Figure 3.38.

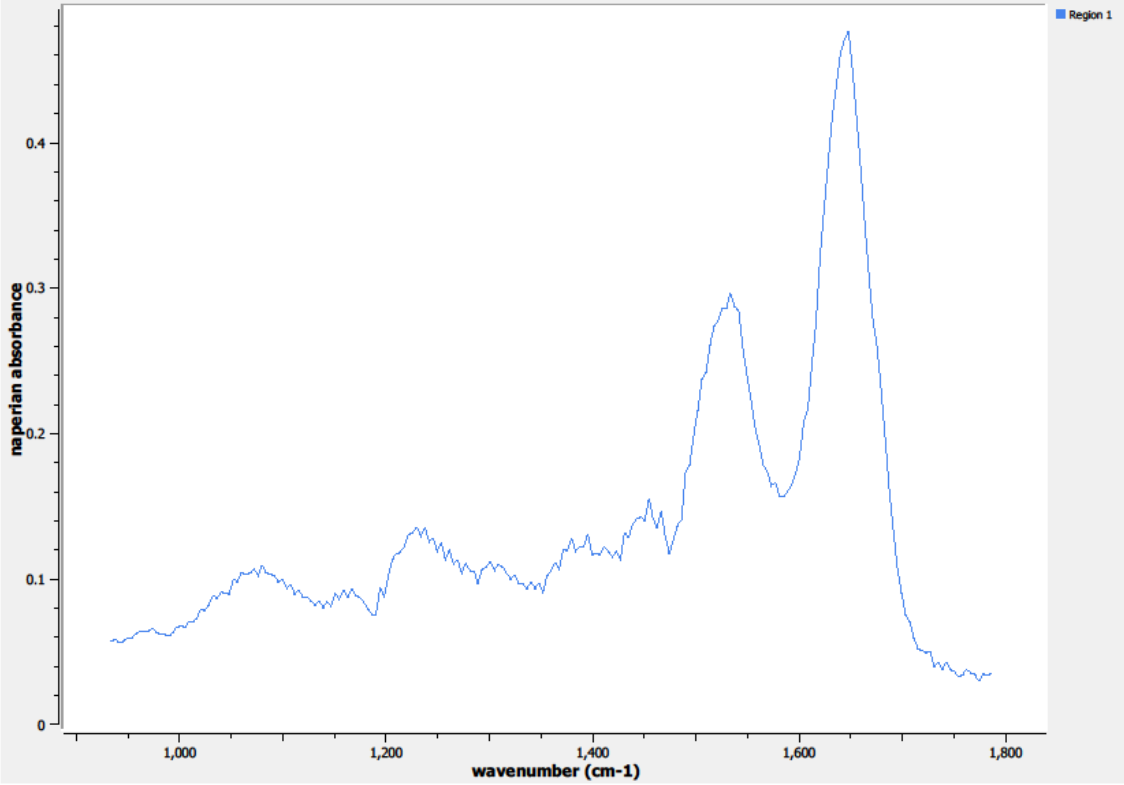



**Figure 3.38. Specimen 13-05-A013h, Point S Spectra**

### 3.3.2.2. Specimen 12-10-A032f

Specimen 12-10-A032f is a malignant breast sample from an 80-year-old white female. The specimen is a 1.33 gram primary malignant breast sample taken from the lobular region. The quality control diagnosis provided by CHTN personnel classifies the specimen as malignant lobular adenocarcinoma of the breast. The tumor percentage for Specimen 12-10-A032f is identified as 90% tumor. The specimen was classified as T2N0M0, which means the tissue is Stage II breast cancer and the tumor is between 2 and 5 centimeters.

The hormonal receptors are diagnosed by CHTN on this specimen as both estrogen positive and progesterone positive. The protein receptor HER2 is identified as negative. An image of a sample from Specimen 12-10-A032f stained with H&E on a glass slide is shown in Figure 3.39. The dark purple throughout the specimen is indicative of an invasive tumor growth permeating throughout the tissue. The aggressiveness of the malignant tissue is evident throughout the entire tissue sample.

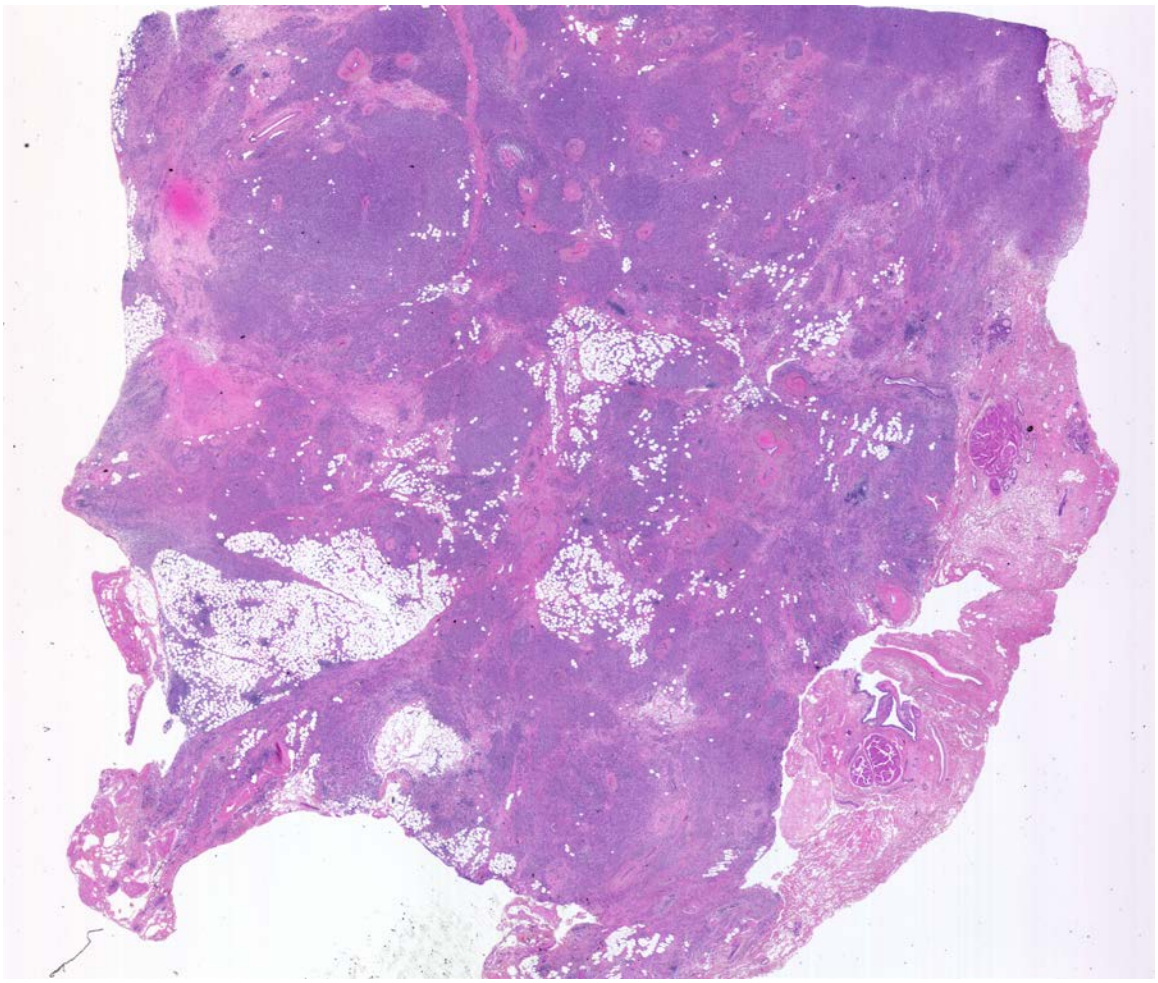


**Figure 3.39. Specimen 12-10-A032f, H&E stain, malignant lobular adenocarcinoma of the breast**

Figure 3.40 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags D, E, and F depict the spectral data collection points on Sample 12-10-A032f, focusing on areas of higher cell concentration.

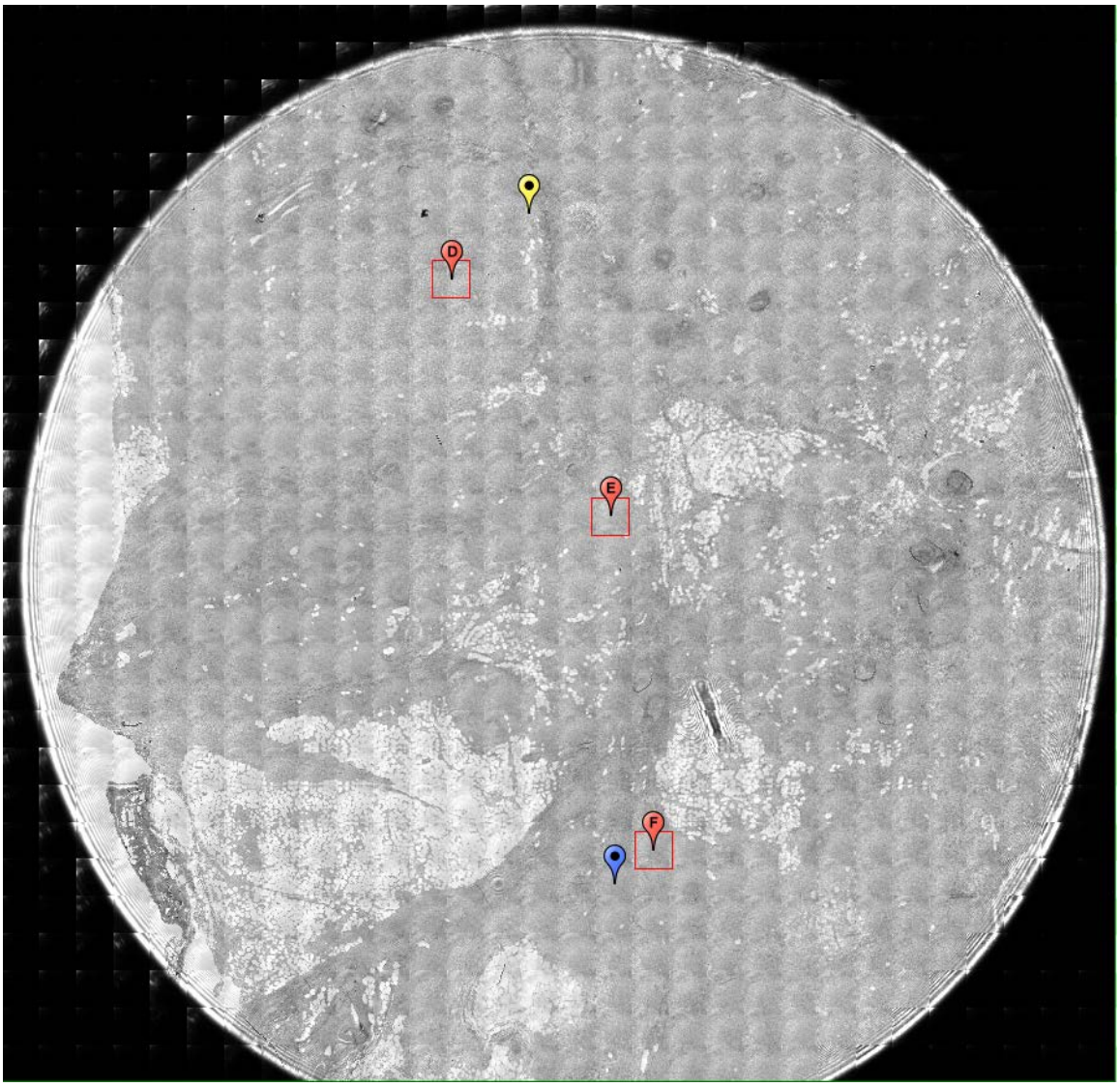



**Figure 3.40. Specimen 12-10-A032f, lobular breast cancer tissue**

#### 3.3.2.2.1. Sample 12-10-A032f, Point D

A 650 x 650 μm field-of-view image taken of the specimen at Point D (see Figure 3.41) was obtained from Spero's high-magnification objective. Point D's location on Sample 12-10-A032f is targeting one of the sample's higher cell

concentration areas, depicted by the dark purple on the H&E stained image. Only one region of interest was needed to capture the spectral information at Point D.

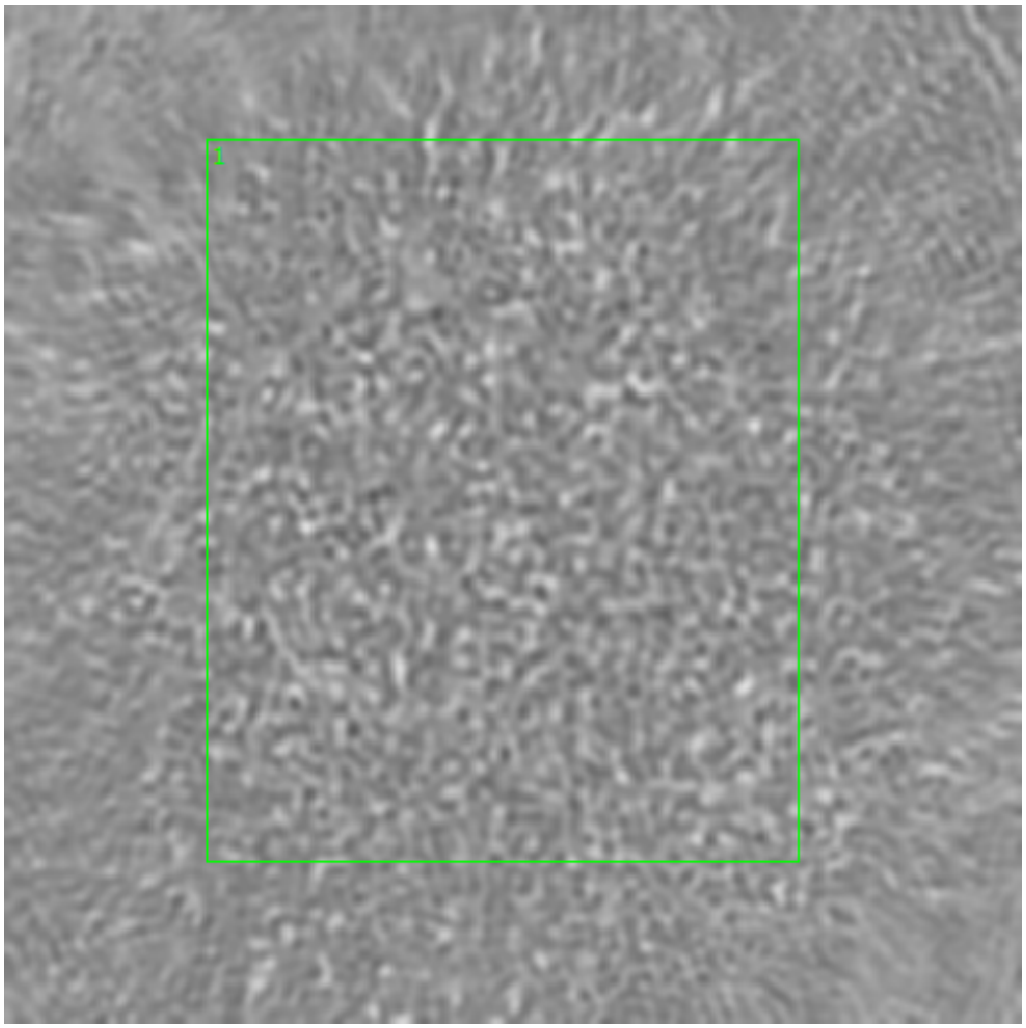



**Figure 3.41. Specimen 12-10-A032f, Point D Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

The spectrum from the one defined region of interest on Point D is depicted in Figure 3.42.

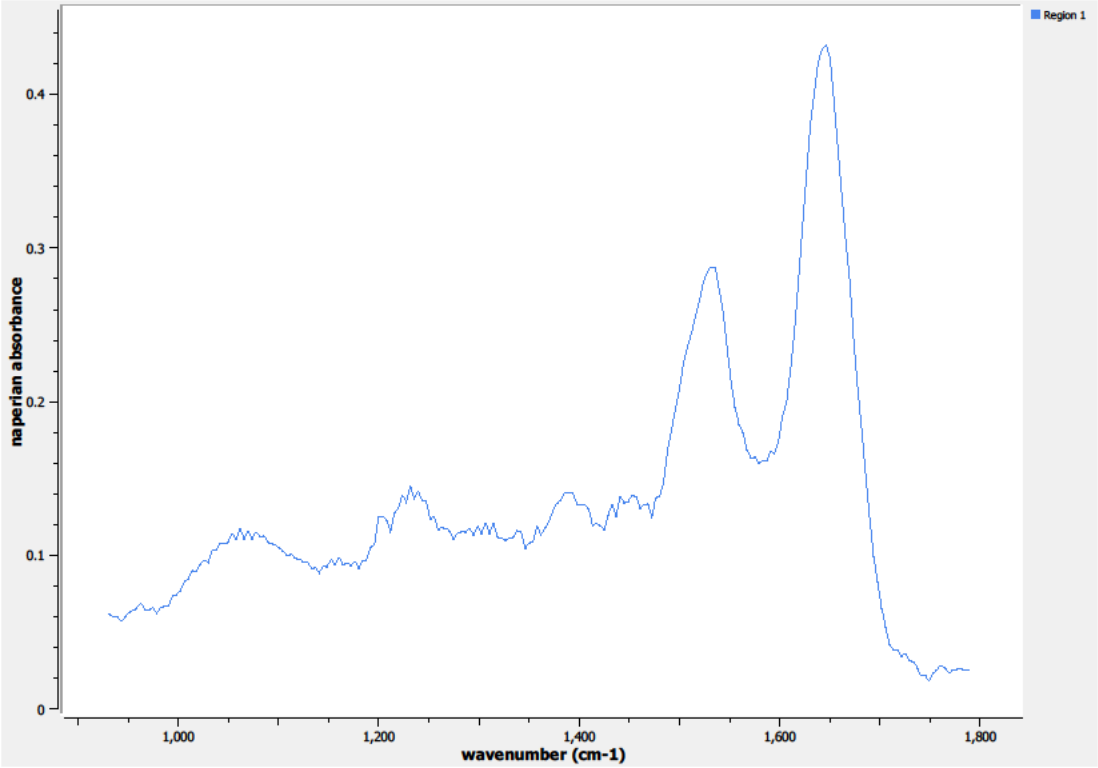



**Figure 3.42. Specimen 12-10-A032f, Point D Spectra**

#### 3.3.2.2.2. Sample 12-10-A032f, Point E

The high-magnification field-of-view image taken at Point E on Sample 12-10-A032f can be seen in Figure 3.43. Point E is again focusing on dense cell concentration areas that indicate a higher tumor percentage.

The spectral data collected on Point E was obtained from the region of interest identified in Figure 3.43. The only region of interest from which spectral information was collected is located on the central portion of the tissue sample.

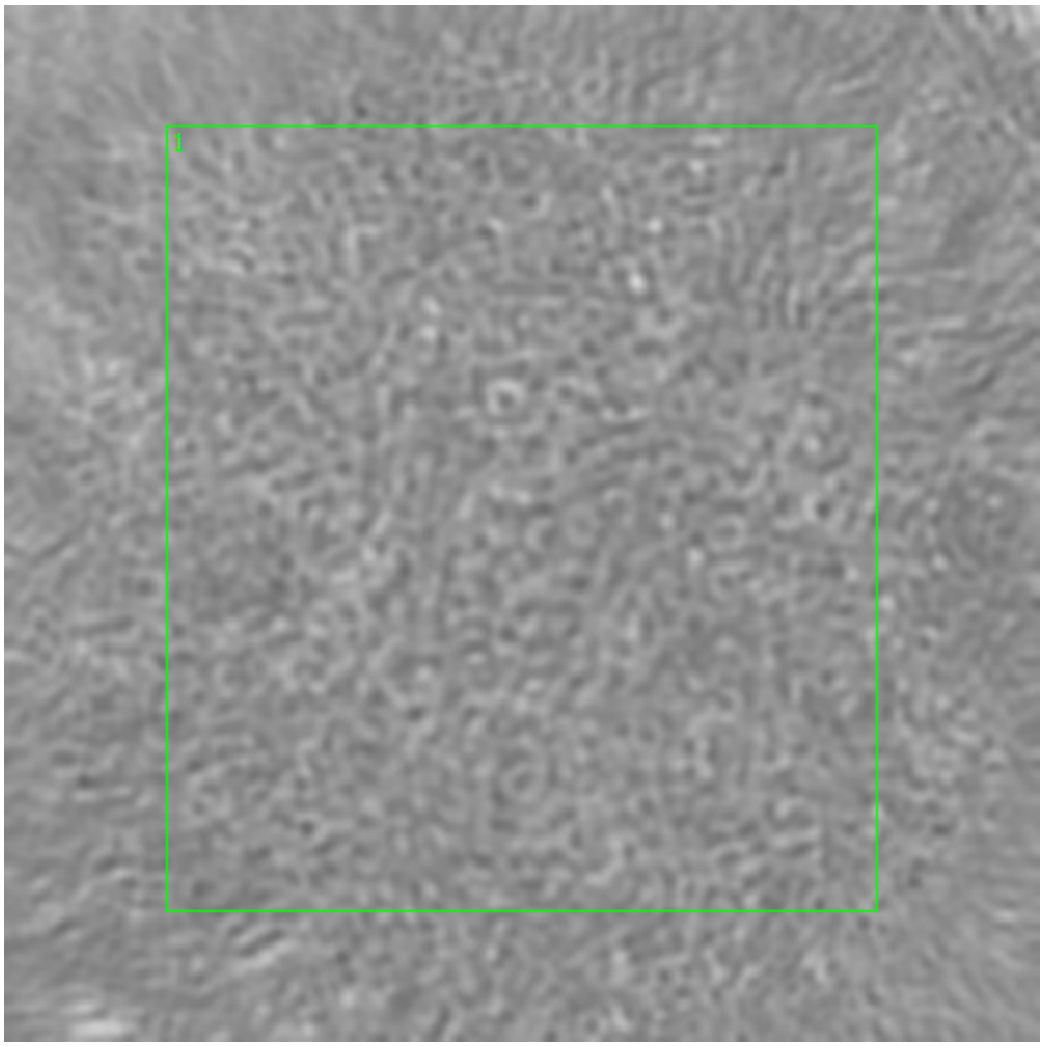



**Figure 3.43. Specimen 12-10-A032f, Point E Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

The spectrum from the one defined region of interest on Point E is shown in Figure 3.44.

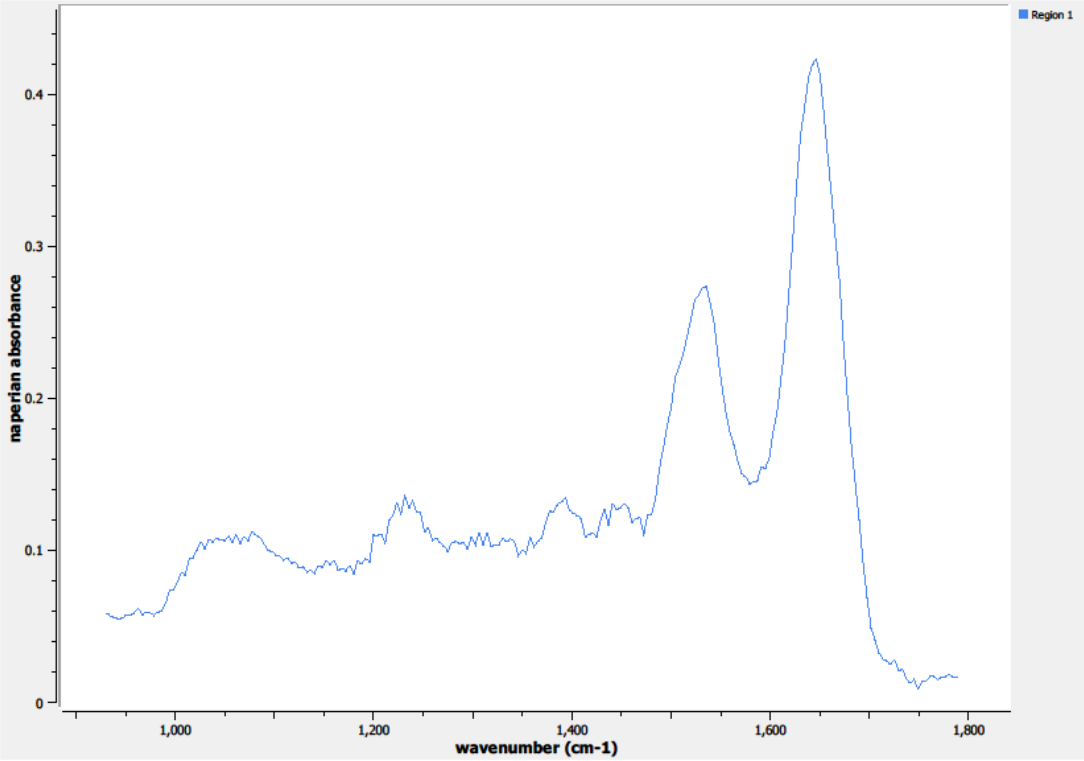



**Figure 3.44. Specimen 12-10-A032f, Point E Spectra**

#### 3.3.2.2.3. Sample 12-10-A032f, Point F

The high-magnification field-of-view image taken at Point F on Sample 12-10-A032f can be seen in Figure 3.45. Point F is again focusing on dense cell concentration areas that indicate a higher tumor percentage. The spectral data collected at Point F was obtained from the single region of interest identified in Figure 3.45.

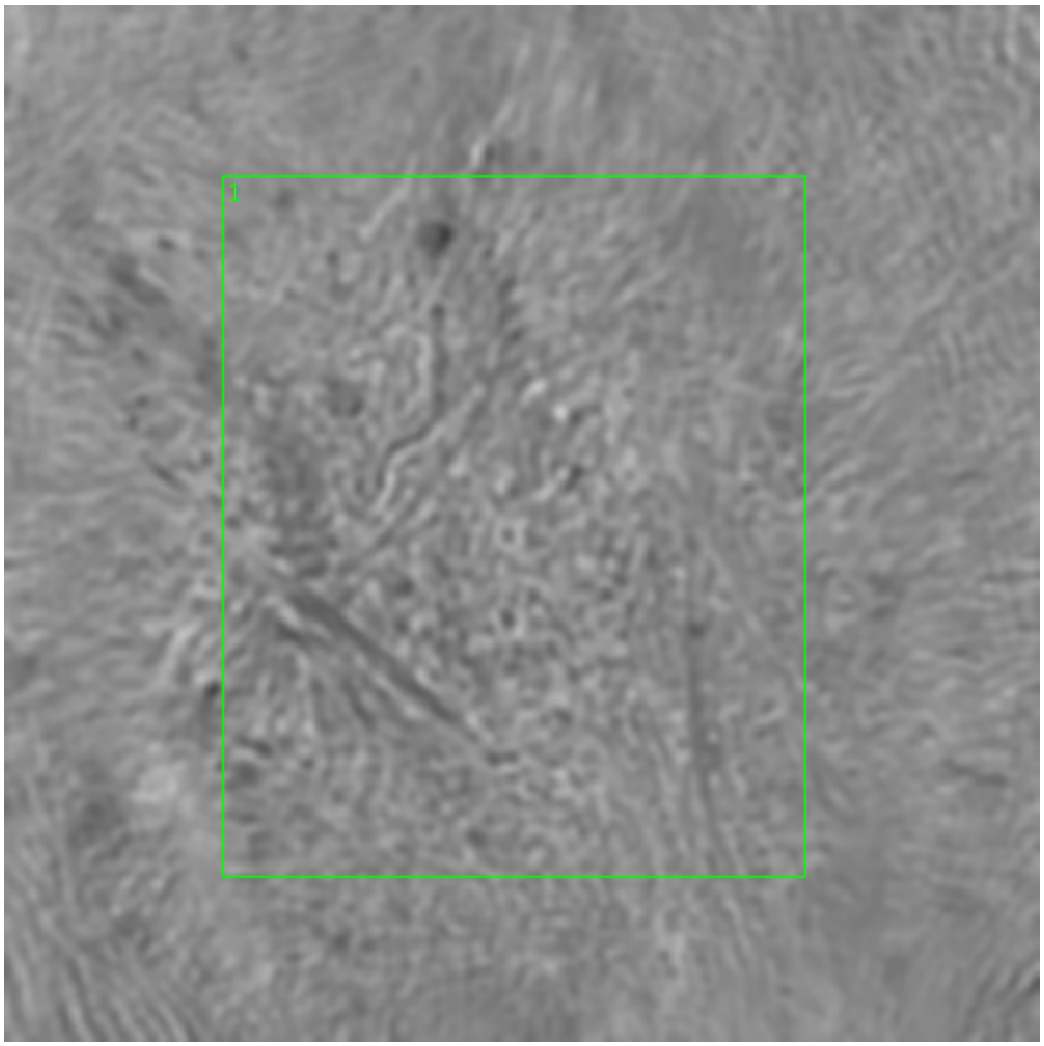



**Figure 3.45. Specimen 12-10-A032f, Point F Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

The spectrum from the one defined region of interest on Point F is shown in Figure 3.46.

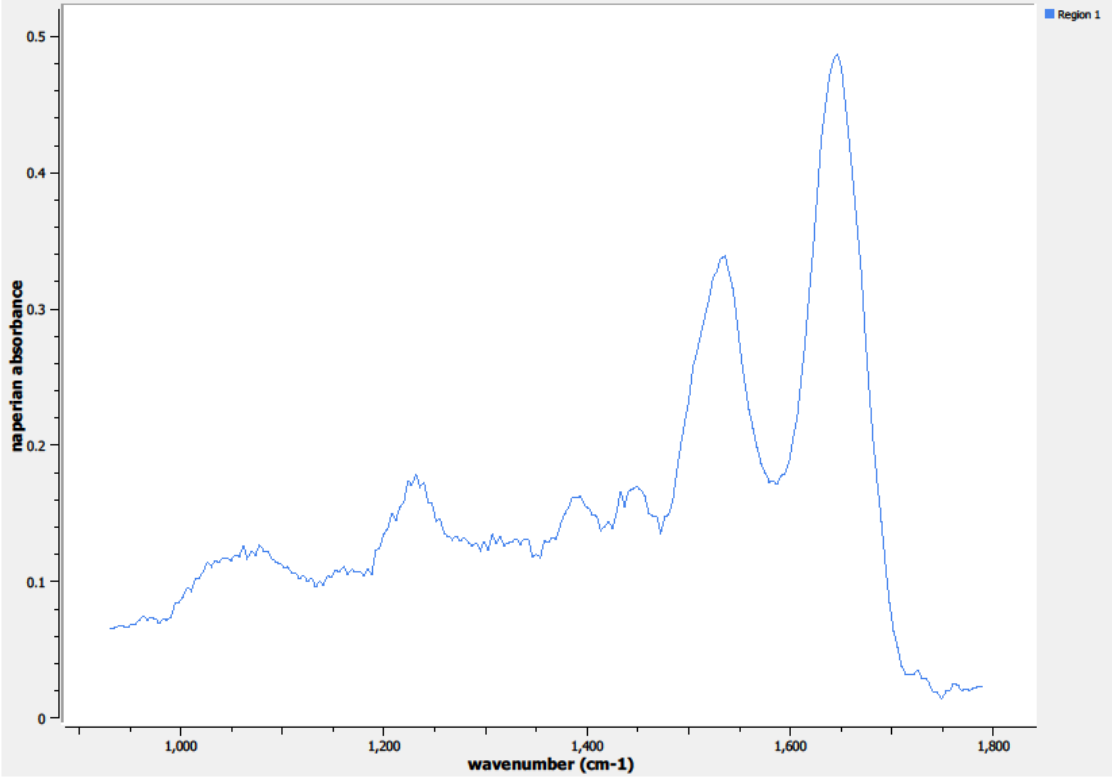



**Figure 3.46. Specimen 12-10-A032f, Point F Spectra**

### 3.3.2.3. Specimen 08-01-A132u

Specimen 08-01-A1332u is a malignant breast sample from a 55-year-old white female. The specimen is a 7.22 gram primary malignant breast sample; this specimen was the densest and largest specimen we received, and the Pathology and Core Laboratory from Baylor College of Medicine had difficulty fitting the incised samples from this specimen on the glass slides and CaF2 windows. CHTN did not include a breast location from which this specimen was removed.

The quality control diagnosis provided by CHTN personnel classifies the specimen as malignant cystosarcoma phylloides of the breast. The tumor percentage for Specimen 08-01-A1332u is identified as 100% tumor. The hormonal receptors and protein receptor diagnoses were not provided by CHTN with this specimen. This specimen is the only example we received that was classified as cystosarcoma phylloides.

Cystosarcoma phylloides tumors are rare and typically benign; the name stems from the leaf-like appearance of the sectioned tissue samples. Phylloides tumors only account for less than 1% of breast tumors, and they are typically large growths in comparison to other breast tumors. Tumors are often misdiagnosed as cystosarcoma phylloides, most likely due to the fleshy characteristics of this type of tumor.[66-69]

Because most cystosarcoma phylloide tumors are benign, the malignant ones are especially dangerous and can lead to delayed diagnosis. Phylloides tumors tend to grow quickly but are generally contained within the breast. Although most cystosarcoma phylloides tumors are benign, our sample was diagnosed as 100% tumor.[67-69]

An image of a sample from Specimen 08-01-A1332u stained with H&E on a glass slide is shown in Figure 3.47. The darker regions of purple on the image indicate higher areas of cell concentration typically associated with a tumor growth. Although the quality control diagnosis was 100% tumor, the H&E sample contains a

fair amount of lighter purple or pink staining, indicating less dense areas of cell concentration. However, even the less dense areas are observably different than the examples of normal breast tissue samples previously listed in Section 3.3.1.

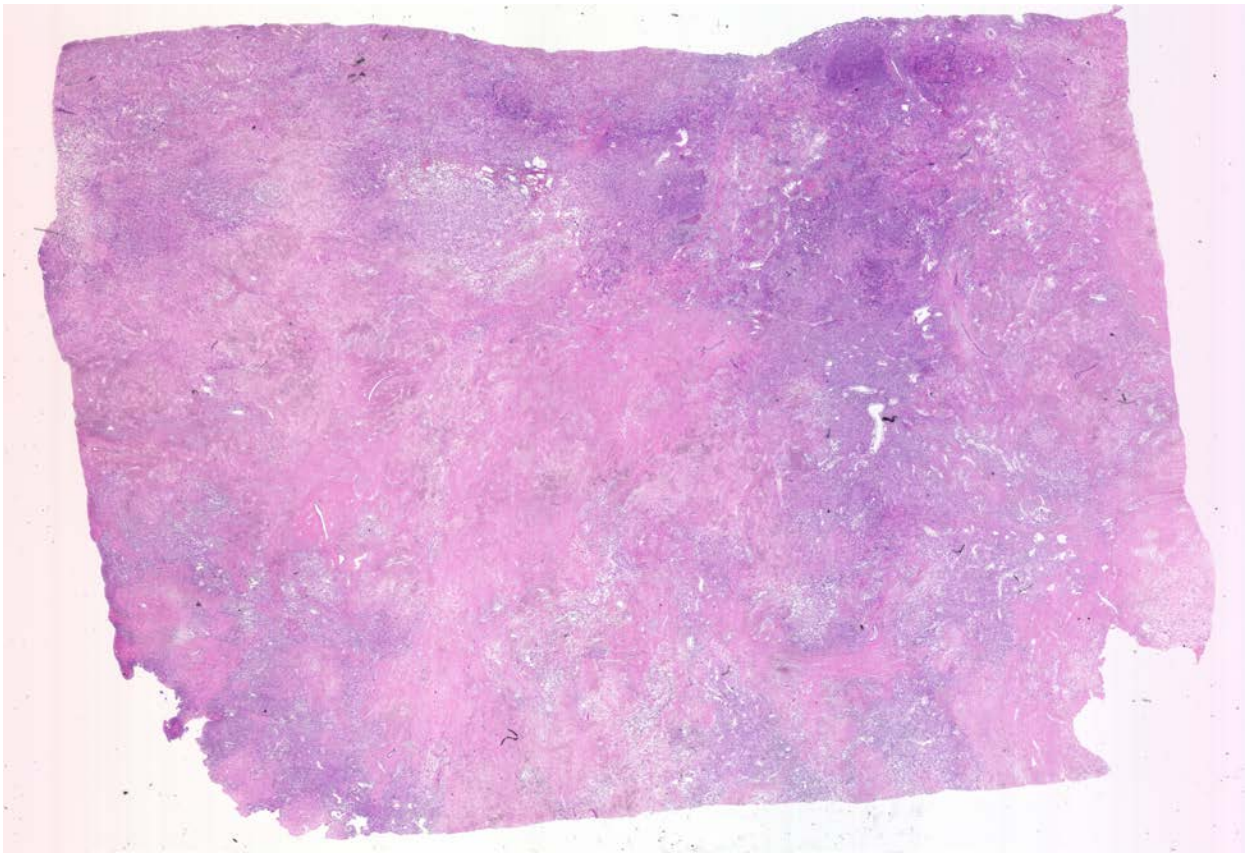


**Figure 3.47. Specimen 08-01-A132u, H&E stain, malignant cystosarcoma phylloides of the breast**

Figure 3.48 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags I and J depict two different spectral data collection points on Sample 08-01-A1332u. Both of these

flags collected spectra in areas of higher cell concentration. Flag K is a third data collection point, although this flag focused on an area of lower cell concentration.

As can be seen in Figure 3.48, the sample size removed from Specimen 08-01-A1332u was physically so large that we had difficulty collecting a background from the remaining portion of the 25 mm diameter CaF2 window void of tissue. Because of the small area that could be utilized as a background, we performed a second background in the same location to ensure unwanted artifacts were not present.

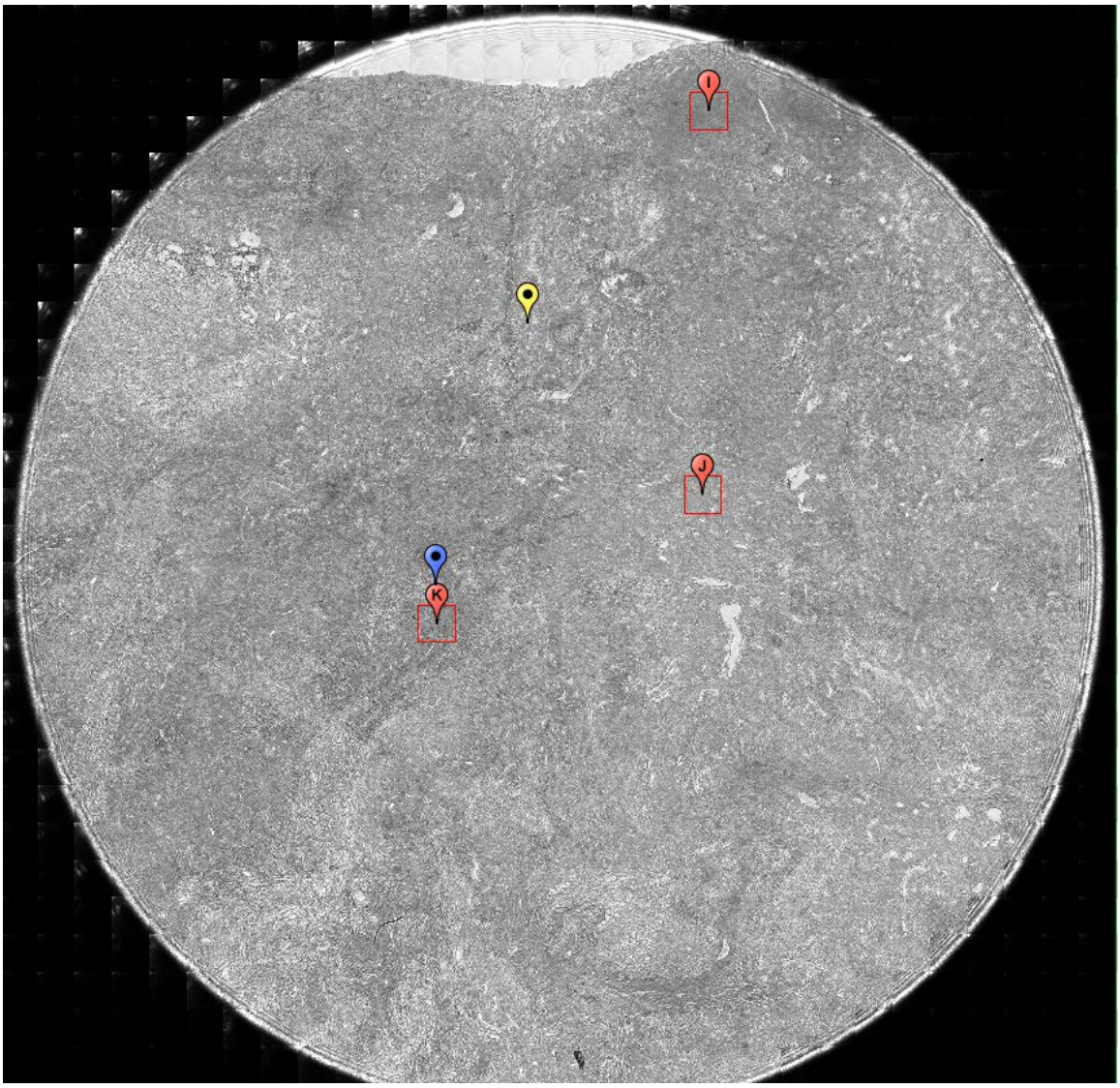



**Figure 3.48. Specimen 08-01-A132u, phylloides breast cancer tissue**

#### 3.3.2.3.1. Sample 08-01-A132u, Point I

Figure 3.49 shows a 650 x 650 μm field-of-view image taken of Sample 08-01-A1332u at Point I. This image was obtained by utilizing Spero's high-magnification objective. Point I is located towards the exterior of Sample 08-01-

A1332u, as shown in Figure 3.48. Point I is targeting one of the sample's areas of higher cell concentration, seen by the dark purple on the H&E stained image. Only one region of interest was needed to capture the spectral information at Point I.

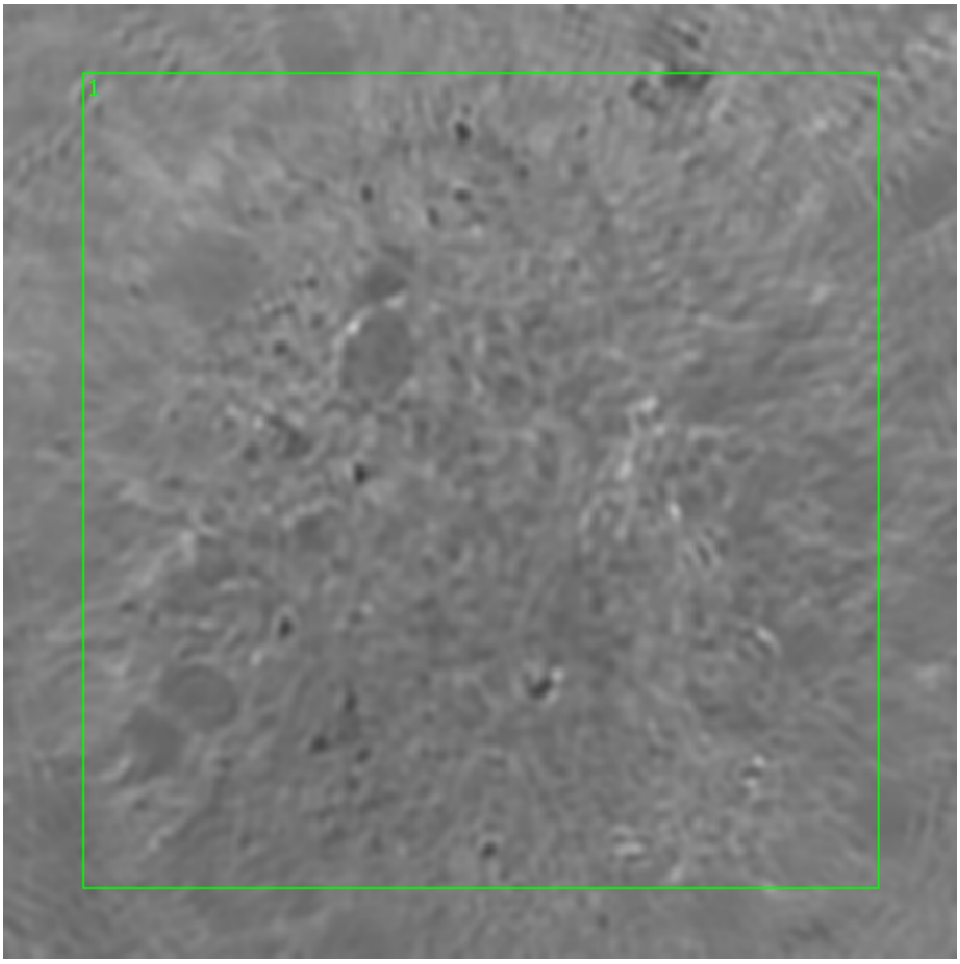



**Figure 3.49. Specimen 08-01-A132u, Point I Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

The spectrum from the defined region of interest on Point I is depicted in Figure 3.50.

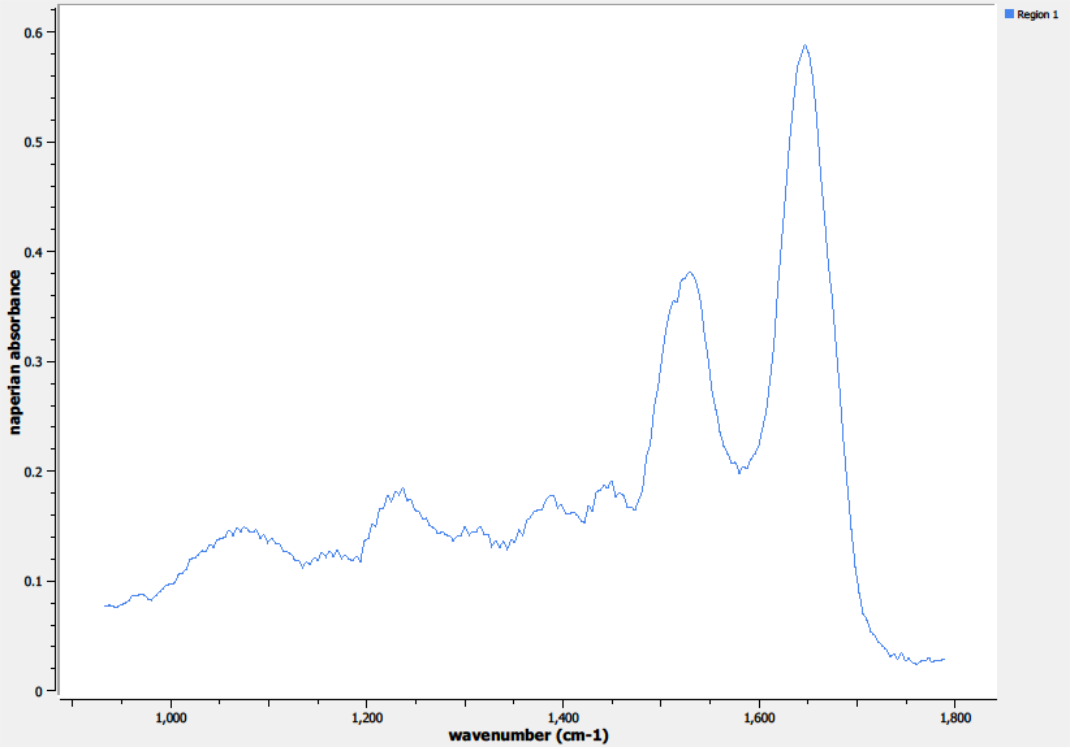



**Figure 3.50. Specimen 08-01-A132u, Point I Spectra**

### 3.3.2.3.2. Sample 08-01-A132u, Point J

A 650 x 650 µm field-of-view image taken of Sample 08-01-A1332u at Point J is seen in Figure 3.51. This image was obtained via Spero's high-magnification objective. Point J is located on the interior of Sample 08-01-A132u; Point J is again focusing on a dense cell concentration area that indicates a higher tumor percentage. Only one region of interest was needed to capture the spectral information at Point J.

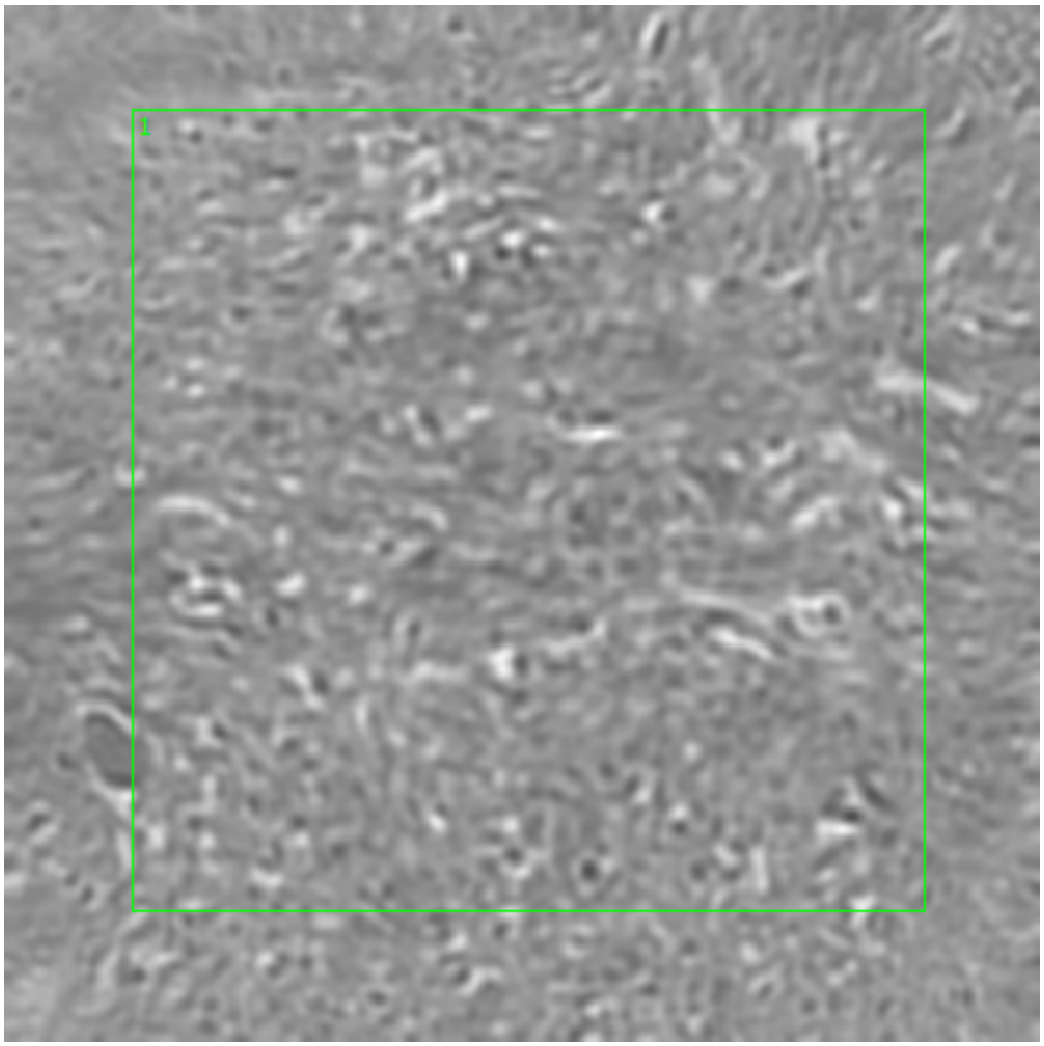



**Figure 3.51. Specimen 08-01-A132u, Point J Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

The spectrum from the defined region of interest on Point J is depicted in Figure 3.52.

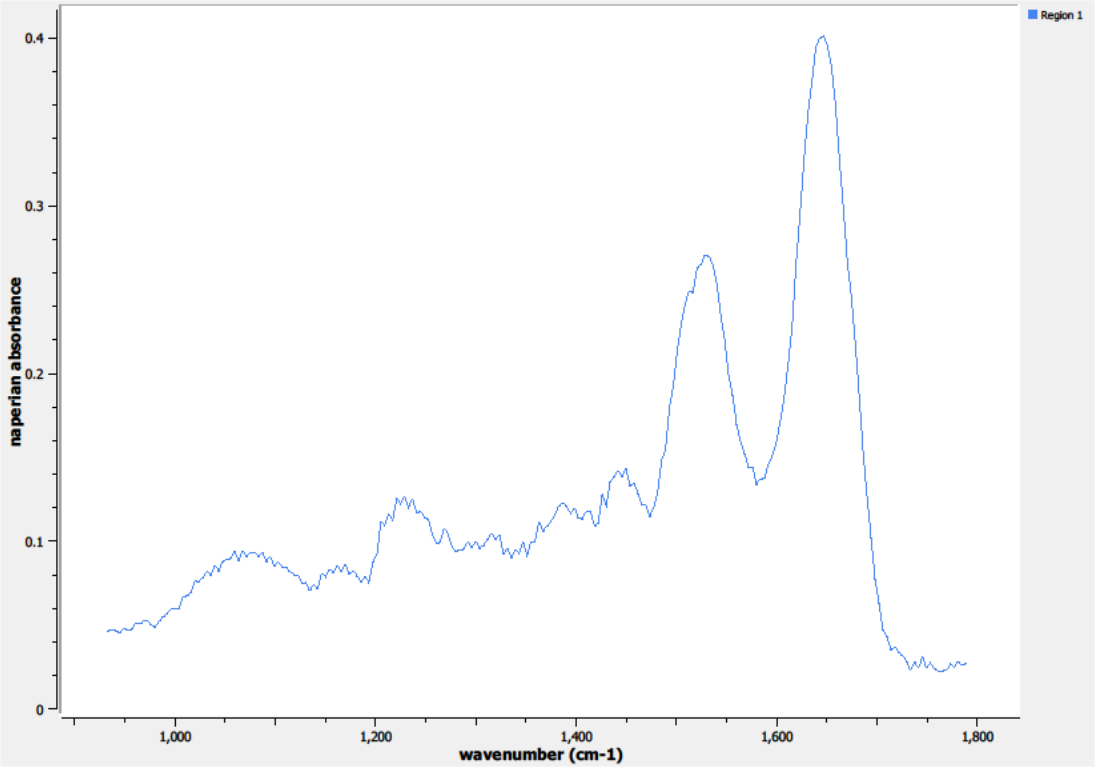



**Figure 3.52. Specimen 08-01-A132u, Point J Spectra**

#### 3.3.2.3.3. Sample 08-01-A132u, Point K

A 650 x 650 µm field-of-view image taken of Sample 08-01-A1332u at Point K is seen in Figure 3.53. This image was obtained via Spero's high-magnification objective. Point K is located on the interior of Sample 08-01-A1332u and is targeting an area that has less cellular concentration than other regions on the sample, which is indicated by the lighter purple or pink coloring of the H&E stained image. Only one region of interest was needed to capture the spectral information at Point J.

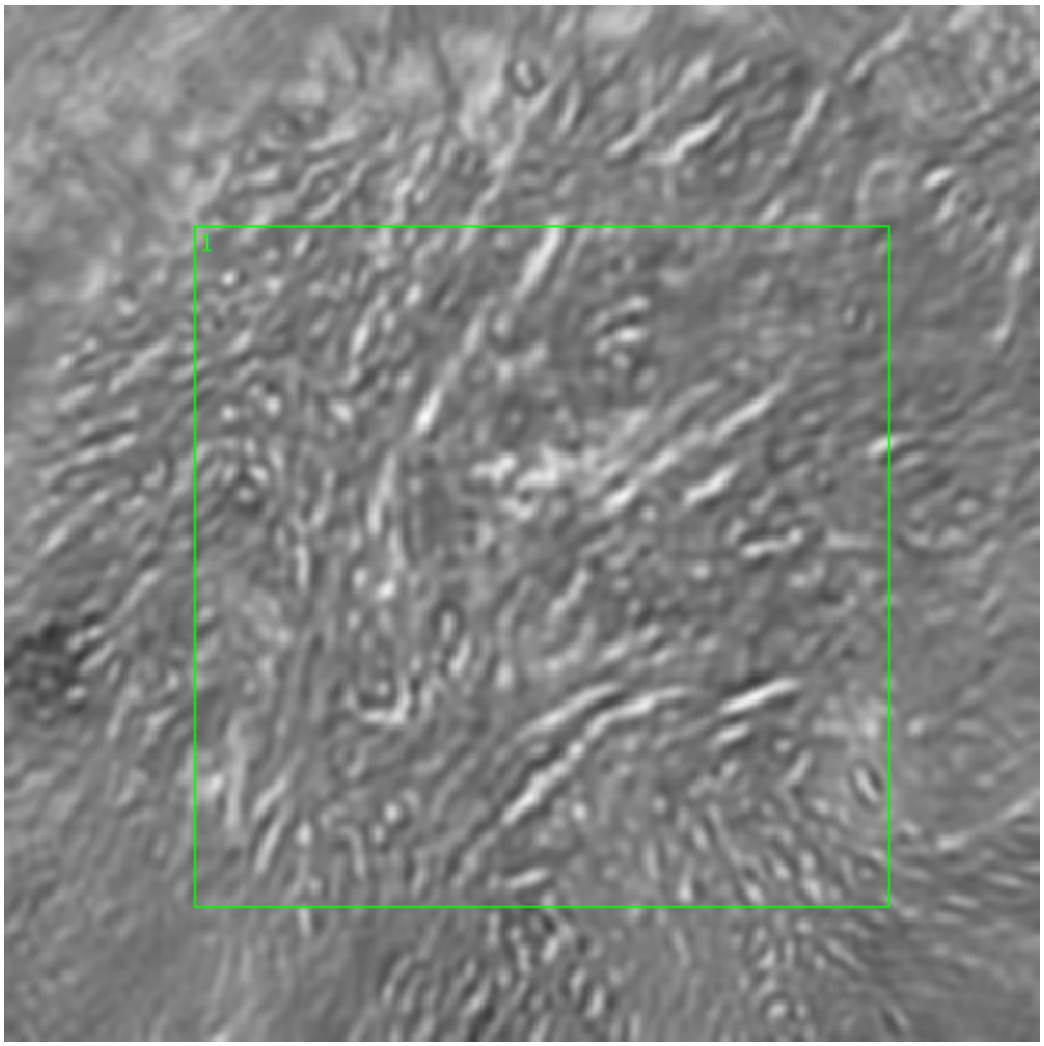



**Figure 3.53. Specimen 08-01-A132u, Point K Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

The spectrum from the defined region of interest on Point K is depicted in Figure 3.54.

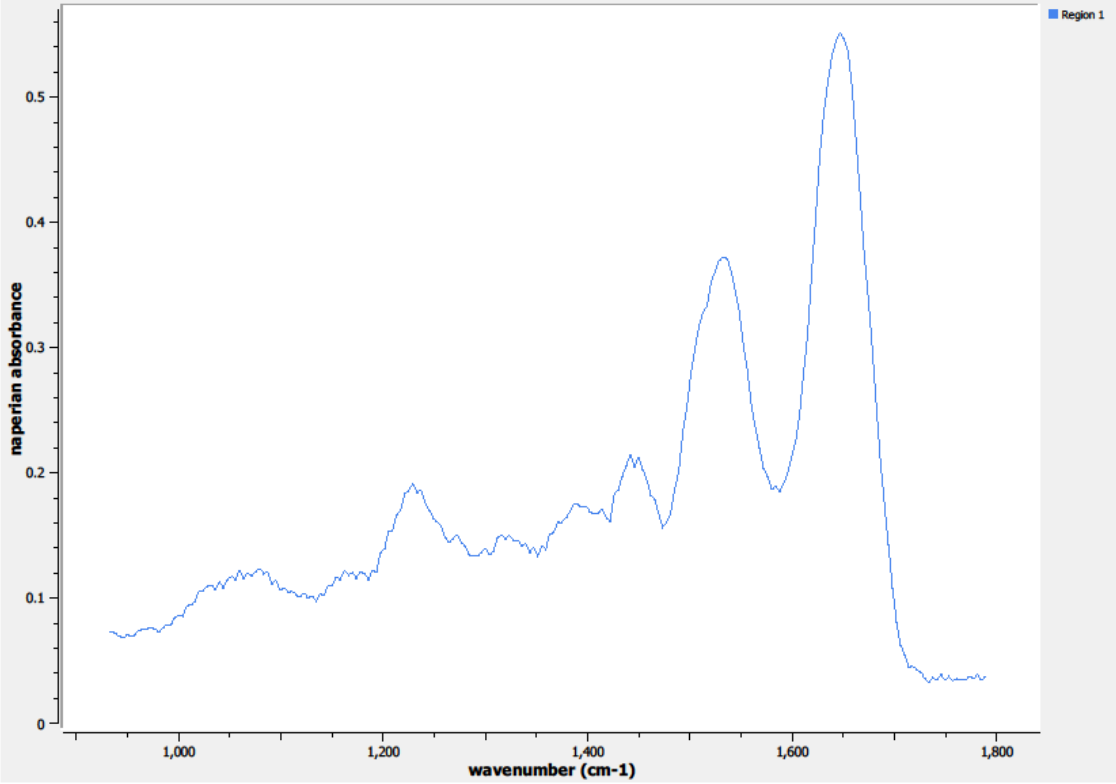



**Figure 3.54. Specimen 08-01-A132u, Point K Spectra**

### 3.3.2.4. Specimen 10-10-A177m

Specimen 10-10-A177m is a malignant breast sample from a 60-year-old white female. The specimen is a 0.68 gram primary malignant breast sample taken from the ductal region. The quality control diagnosis provided by CHTN personnel classifies the specimen as malignant ductal adenocarcinoma of the breast. The tumor percentage for Specimen 10-10-A177m is identified as 100% tumor.

Both hormonal receptors are diagnosed by CHTN on this specimen as estrogen negative and progesterone negative. The protein receptor HER2 is also

identified as negative, making this specimen a “triple negative” diagnosis. This specimen is the only one we received that was classified as “triple negative.”

An image of a sample from Specimen 10-10-A177m stained with H&E on a glass slide is shown in Figure 3.55. The darker regions of purple on the image indicate higher areas of cell concentration typically associated with tumor growth. The invasiveness of the malignant tumor growth is evident throughout the entire tissue sample.

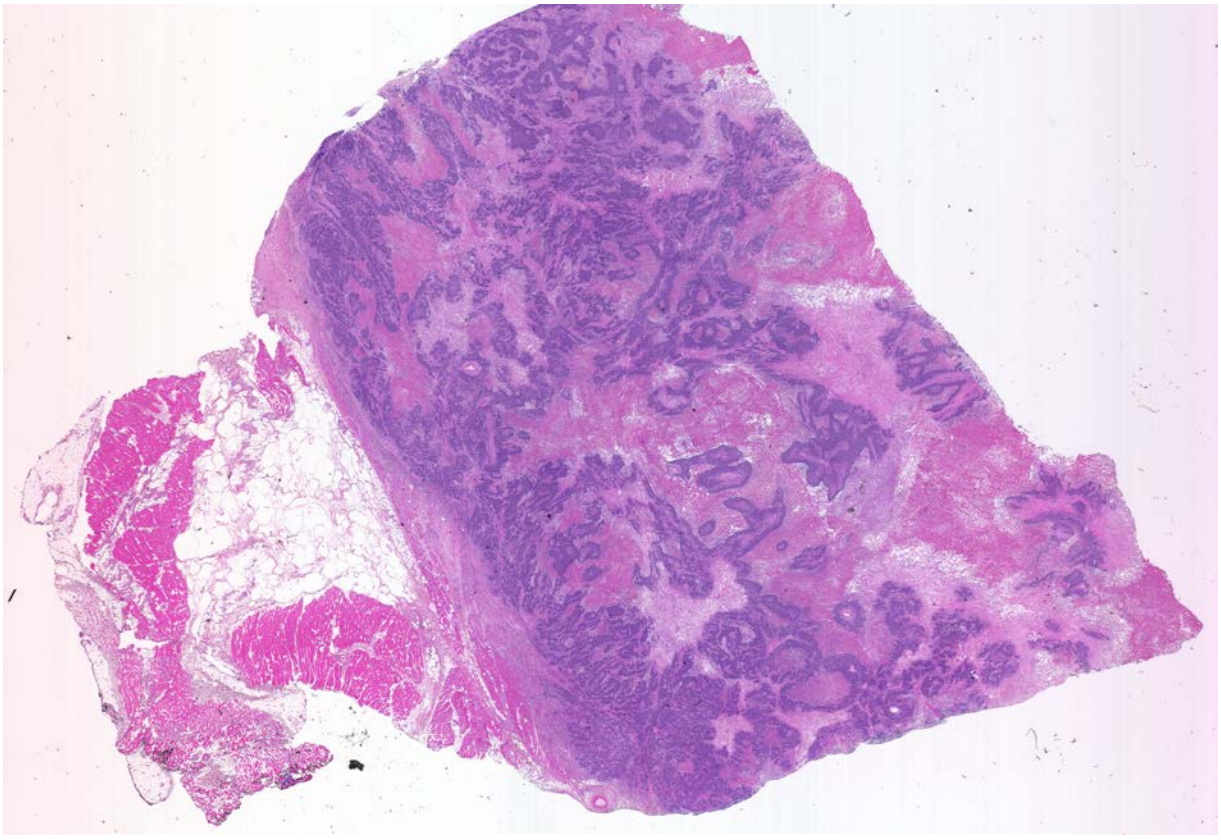


**Figure 3.55. Specimen 10-10-A177m, H&E stain, malignant ductal adenocarcinoma of the breast, triple negative**

Triple negative tumors are a rare and deadly form of breast cancer. Triple negative tumors account for approximately 10% of all malignant breast tumors. These tumors are often quite aggressive and commonly recur in patients. Patients with triple negative tumors often do not survive due to late diagnoses.[70-78]

Because triple negative tumors do not have a readily identifiable molecular target, they can be very difficult to identify due to the void of the common biomarkers. This tumor type typically does not respond to receptor-targeted or hormonal treatment programs, which are some of the more common treatment programs for breast cancer. In regards to both identification and treatment, it is difficult to target something that does not exist.[70-78]

Specimen 10-10-A177m is the most interesting malignant breast sample we have due to the triple negative diagnosis. As previously stated, early detection is one of the greatest defenses against cancer. Triple negative tumors are one of the more dangerous forms of breast cancer due to the negative readings on all three main receptors, so they often goes undiagnosed in the early stages of cancer.

According to Dr. Sharma from the University of Kansas Cancer Center, 50-70% of triple negative tumors have DNA mutations, specifically the BRCA gene. Mid-IR detection could provide a unique analysis technique to find triple negative tumors by being able to identify these DNA mutations using spectroscopy.[75-77]

Figure 3.56 is an image obtained from Spero using the high-magnification objective and the mapping feature of the instrument. Flags U, V, W, and X depict

four different spectral data collection points on Sample 10-10-A177m. Flag U identifies a region of high cell concentration; Flags W and X are also targeting high cellular concentration regions within the tumor, but the concentration is not as high as at Flag U. Flag V pinpoints an area that may not be part of the malignant tumor; the H&E stain seen in Figure 3.55 shows this region as being a deep pink coloring, which, although it is not what is typically seen on the tumor portion of a malignant breast tissue sample, indicates some abnormality while not necessarily being part of the malignant tumor growth.

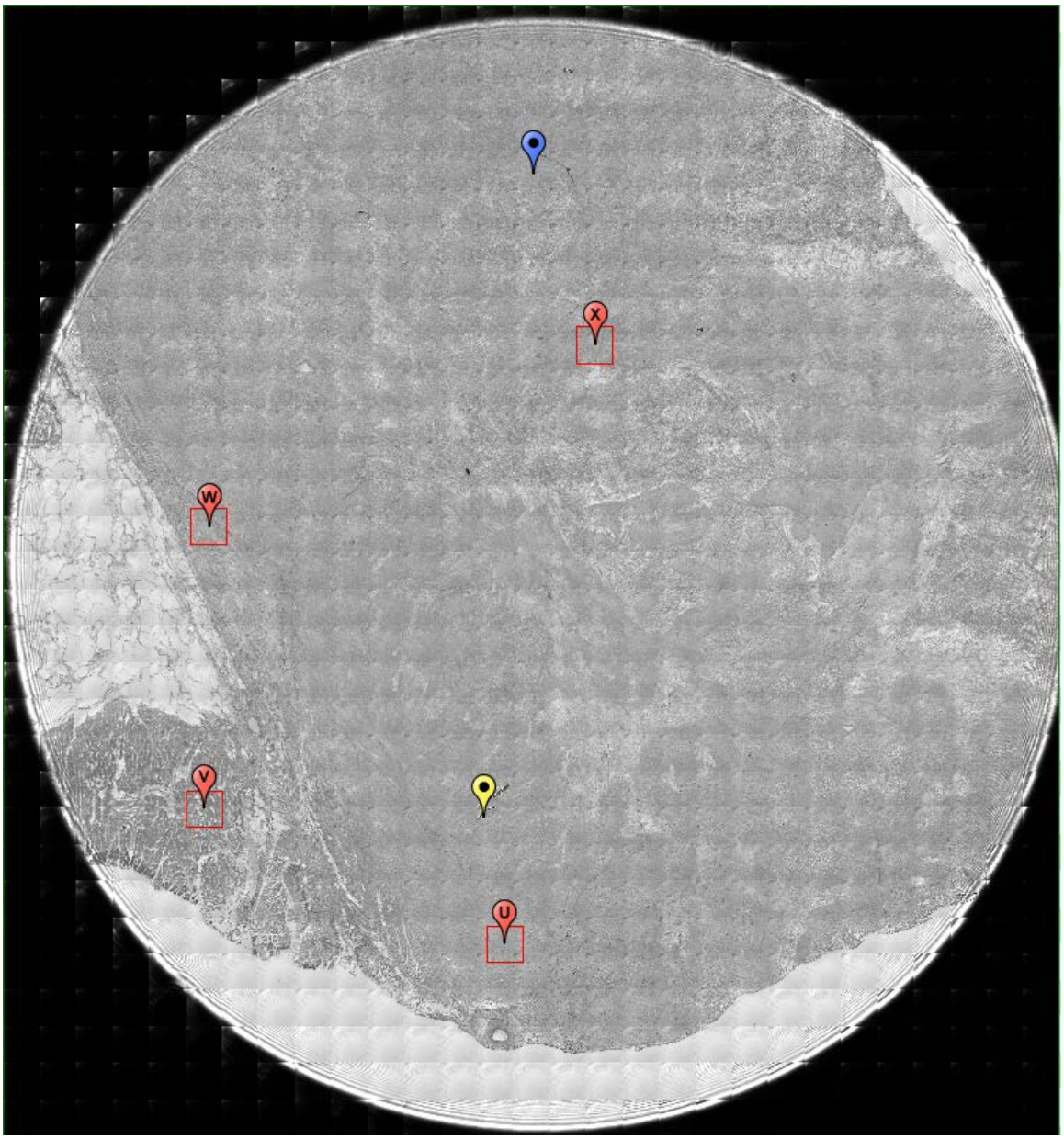



**Figure 3.56. Specimen 10-10-A177m, ductal breast cancer tissue, triple negative**

#### 3.3.2.4.1. Sample 10-10-A177m, Point U

A 650 x 650 µm field-of-view image taken of Sample 10-10-A177m at Point U (see Figure 3.57) was obtained from Spero's high-magnification objective. Point U is

located in the midst of one of the densest areas of the malignant cell growth on this sample, depicted by the dark purple on the H&E stained image. Only one region of interest was needed to capture the spectral information at Point U.

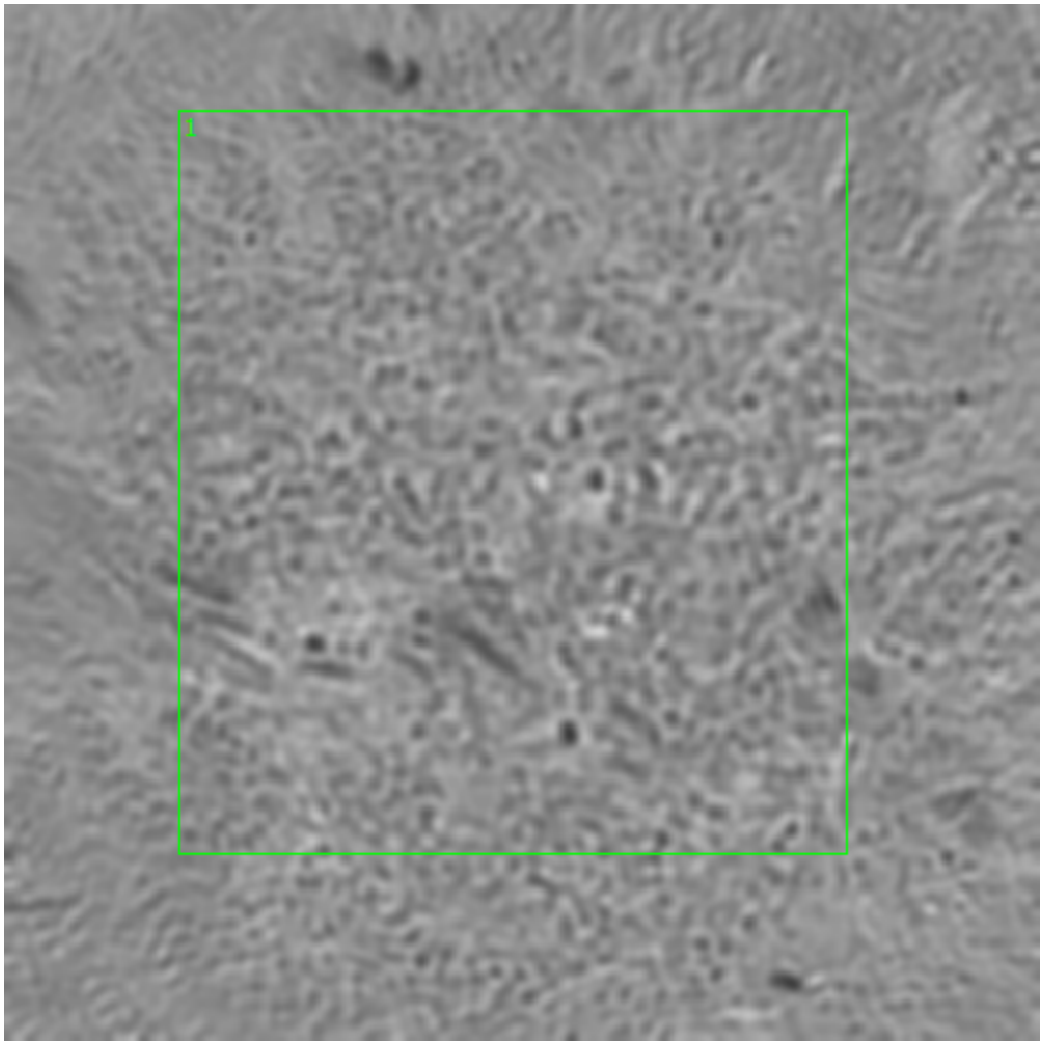



**Figure 3.57. Specimen 10-10-A177m, Point U Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative**

The spectrum from the defined region of interest on Point U is depicted in Figure 3.58.

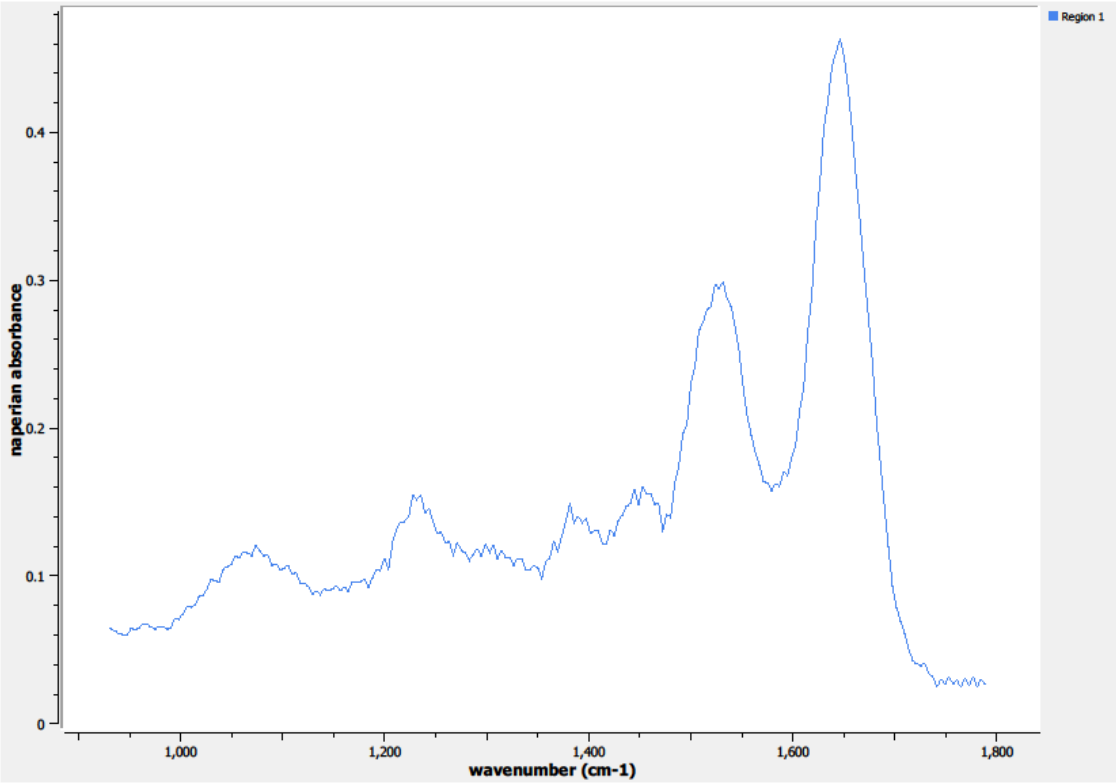



**Figure 3.58. Specimen 10-10-A177m, Point U Spectra**

#### 3.3.2.4.2. Sample 10-10-A177m, Point V

The high-magnification field-of-view image taken at Point V on Sample 10-10-A177m can be seen in Figure 3.59. Point V's focus is on the portion of tissue that does not appear to be part of the malignant growth seen with most typical breast tumors on the H & E-stained image; whether or not this portion of the tissue is

malignant, though, by observation that portion of the sample's tissue appears to be different than the normal breast tissue seen from the previous section.

The spectral data collected at Point V was obtained from the regions of interest identified in Figure 3.59. Region 1 is a general region of interest over the center of the field-of-view; Regions 2, 3, 4, 5, 6, and 7 are defining denser tissue areas for spectral data collection. Again, by observation, even the high-magnification image seen in Figure 3.59 appears to be significantly different than the high-magnification images from the normal breast specimens presented in Section 3.3.1.

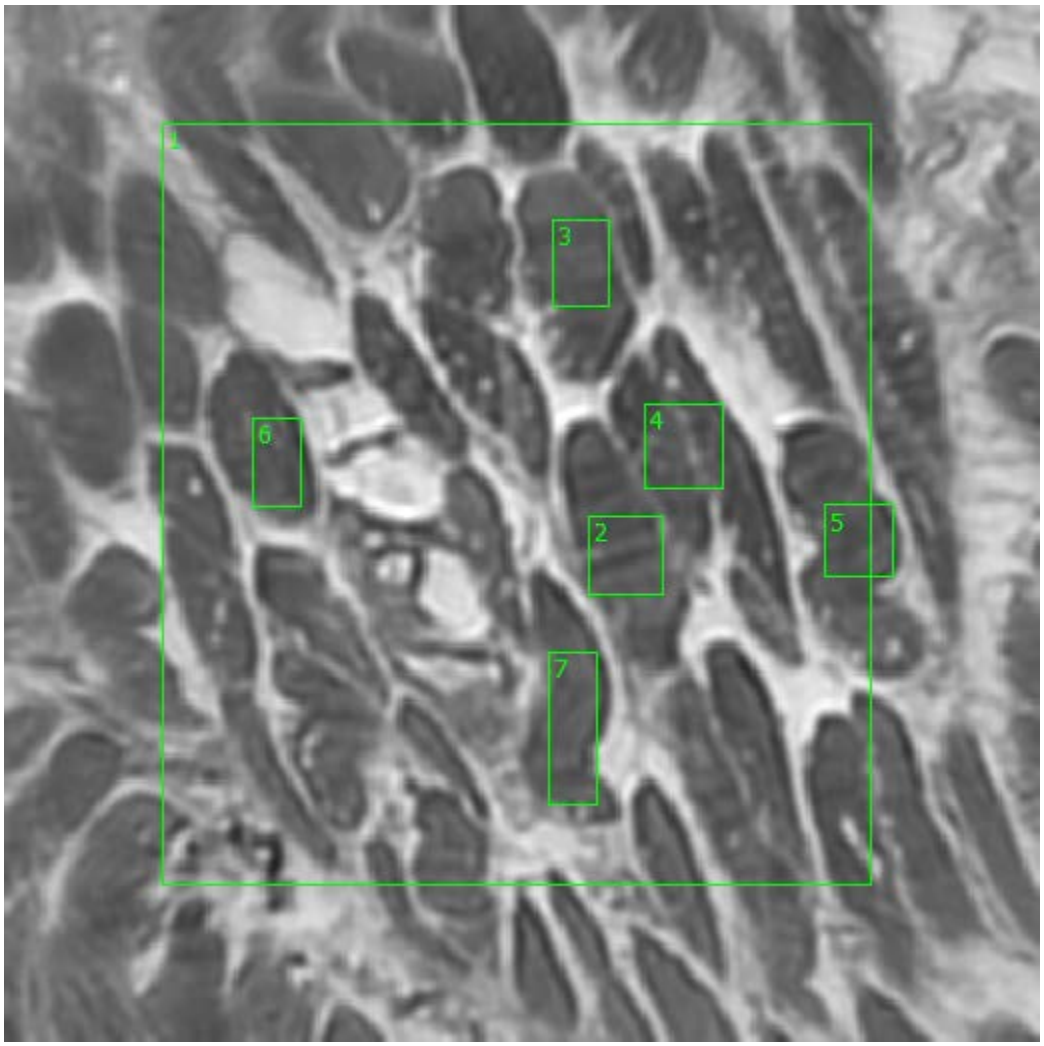



**Figure 3.59. Specimen 10-10-A177m, Point V Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative, transition region**
**Regions 2, 3, 4, 5, 6, 7: breast, malignant ductal adenocarcinoma, triple negative, condensed transition region**

The spectra from the defined regions of interest on Point V are depicted in Figure 3.60.

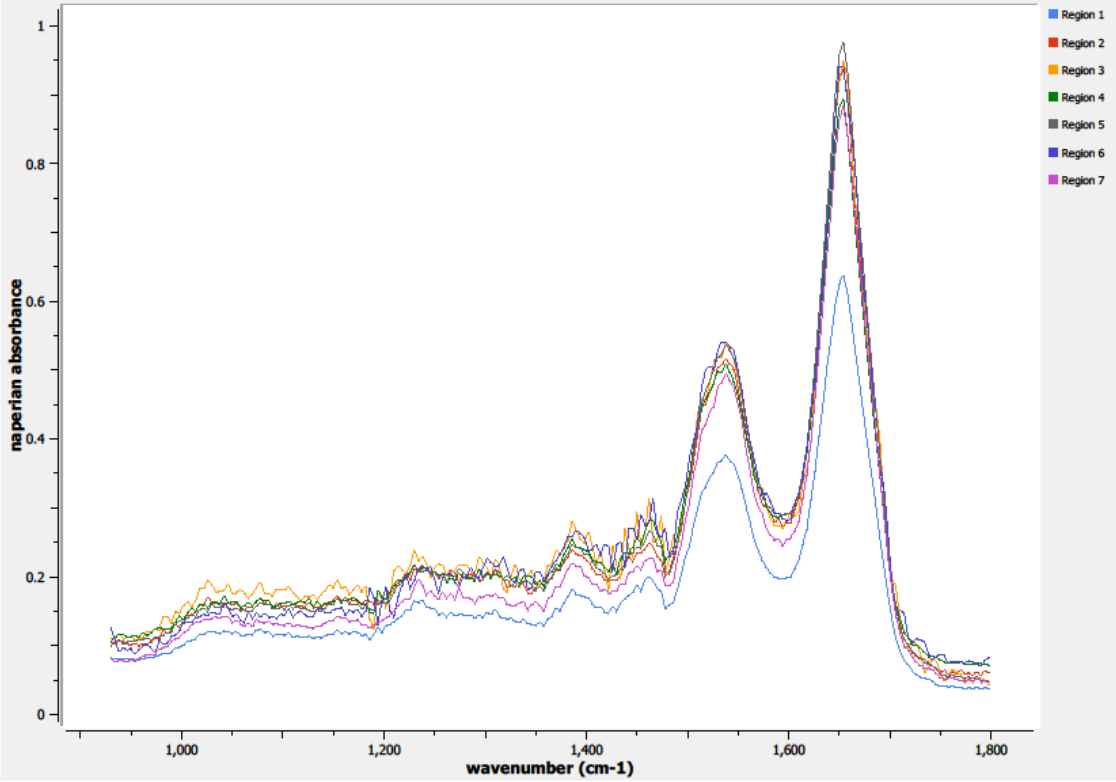



**Figure 3.60. Specimen 10-10-A177m, Point V Spectra**

#### 3.3.2.4.3. Sample 10-10-A177m, Point W

A 650 x 650 µm field-of-view image taken of Sample 10-10-A177m at Point W (see Figure 3.61) was obtained from Spero's high-magnification objective. Point W is located in the midst of one of the densest areas of the malignant cell growth on this sample, depicted by the dark purple on the H&E stained image. Two regions of interest were obtained to capture the spectral information at Point W. Region 1 may not be quite as dense as Region 2, according to the H & E stain, although the total

absorption seen in the spectra in Figure 3.62 is slightly higher for Region 1 than Region 2.

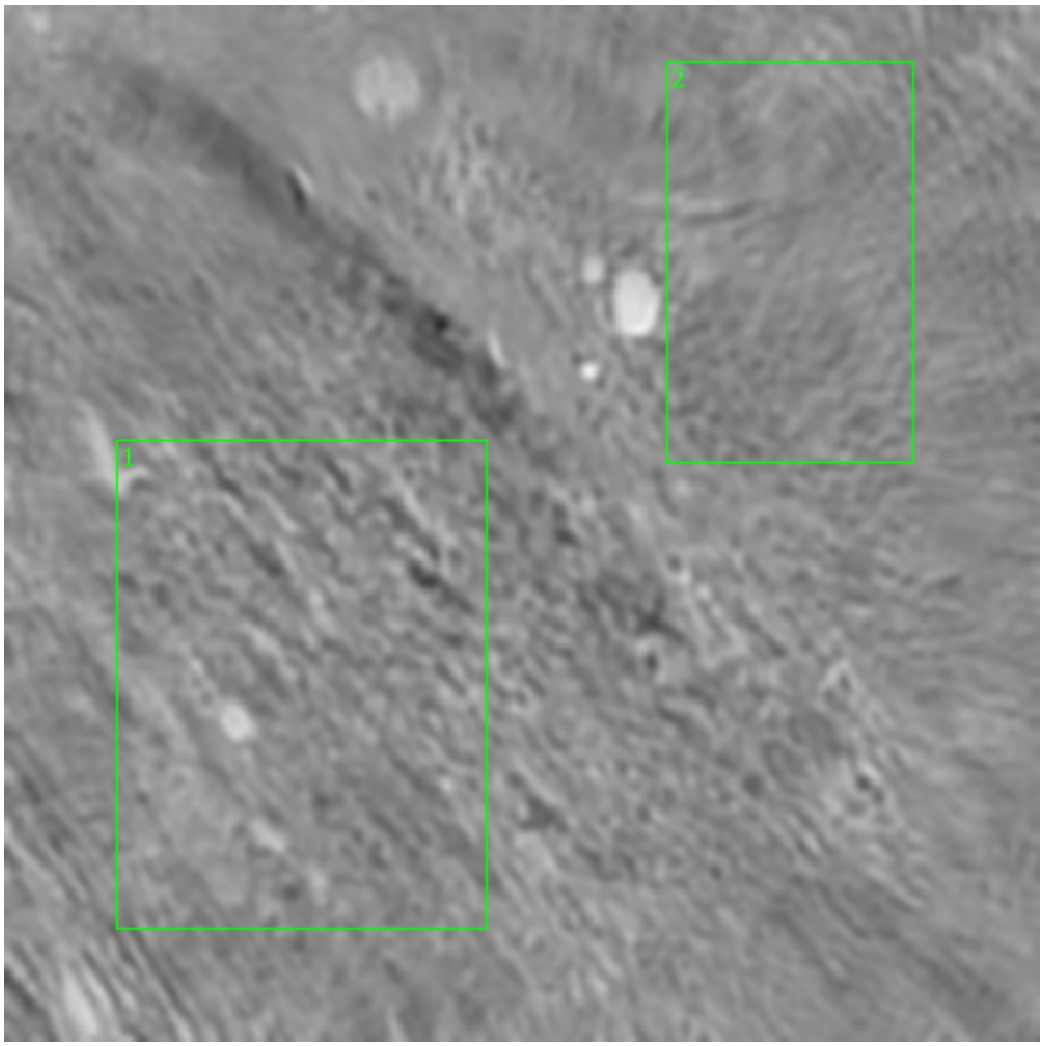



**Figure 3.61. Specimen 10-10-A177m, Point W Regions of Interest**
**Regions 1, 2: breast, malignant ductal adenocarcinoma, triple negative**

The spectrum from the defined region of interest on Point W is depicted in Figure 3.62.

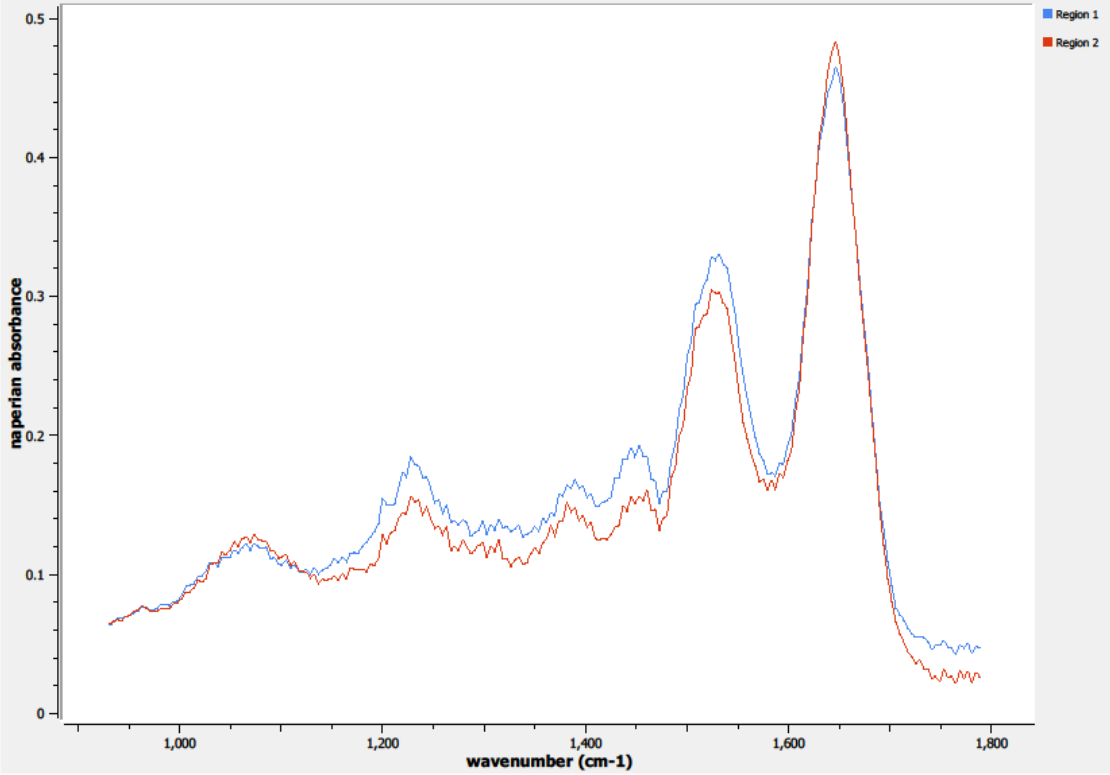



**Figure 3.62. Specimen 10-10-A177m, Point W Spectra**

#### 3.3.2.4.1. Sample 10-10-A177m, Point X

A 650 x 650 µm field-of-view image taken of Sample 10-10-A177m at Point X is seen in Figure 3.63. This image was obtained via Spero's high-magnification objective. Point X is located on the interior of Sample 10-10-A177m; Point X is again focusing on a dense cell concentration area that indicates a higher tumor percentage. Only one region of interest was needed to capture the spectral information at Point X.

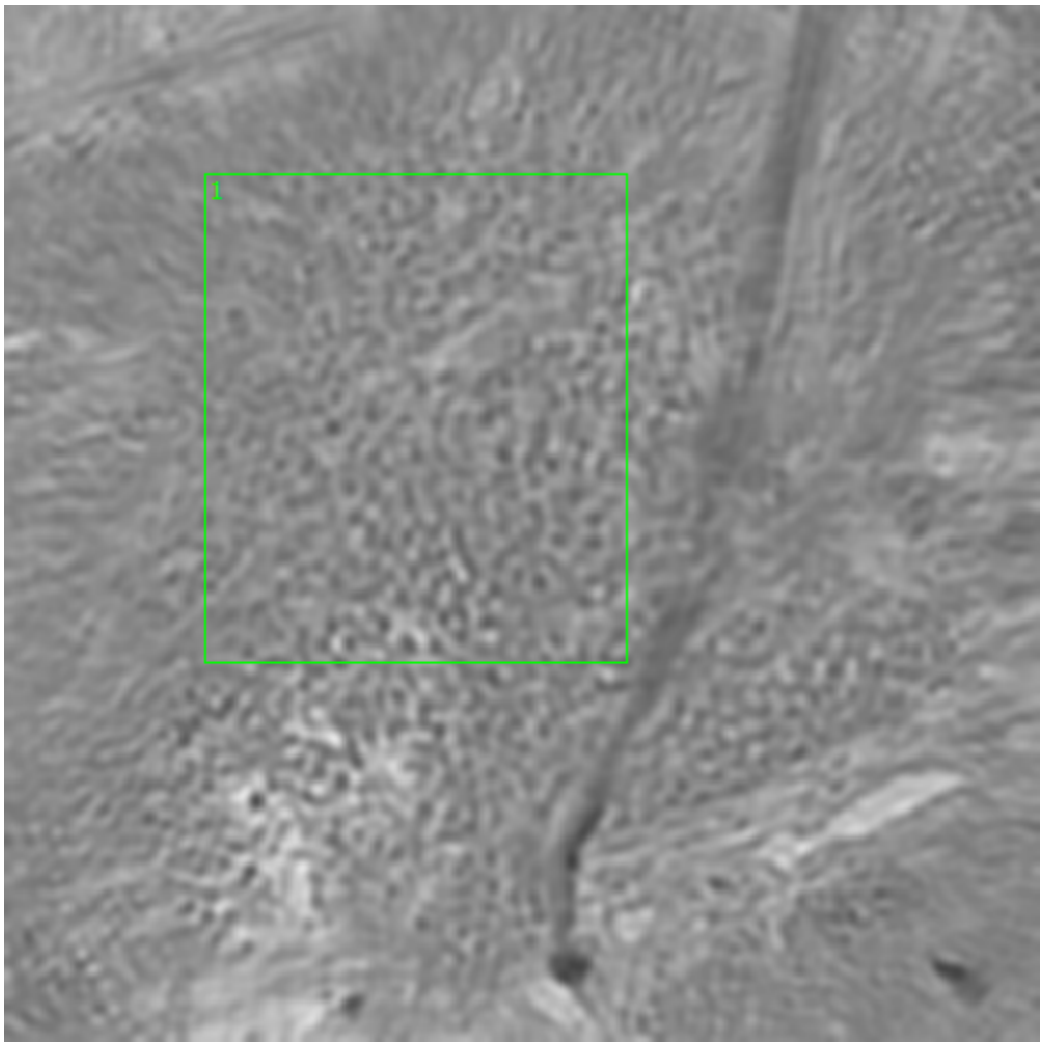



**Figure 3.63. Specimen 10-10-A177m, Point X Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative**

The spectrum from the defined region of interest on Point X is depicted in Figure 3.64.

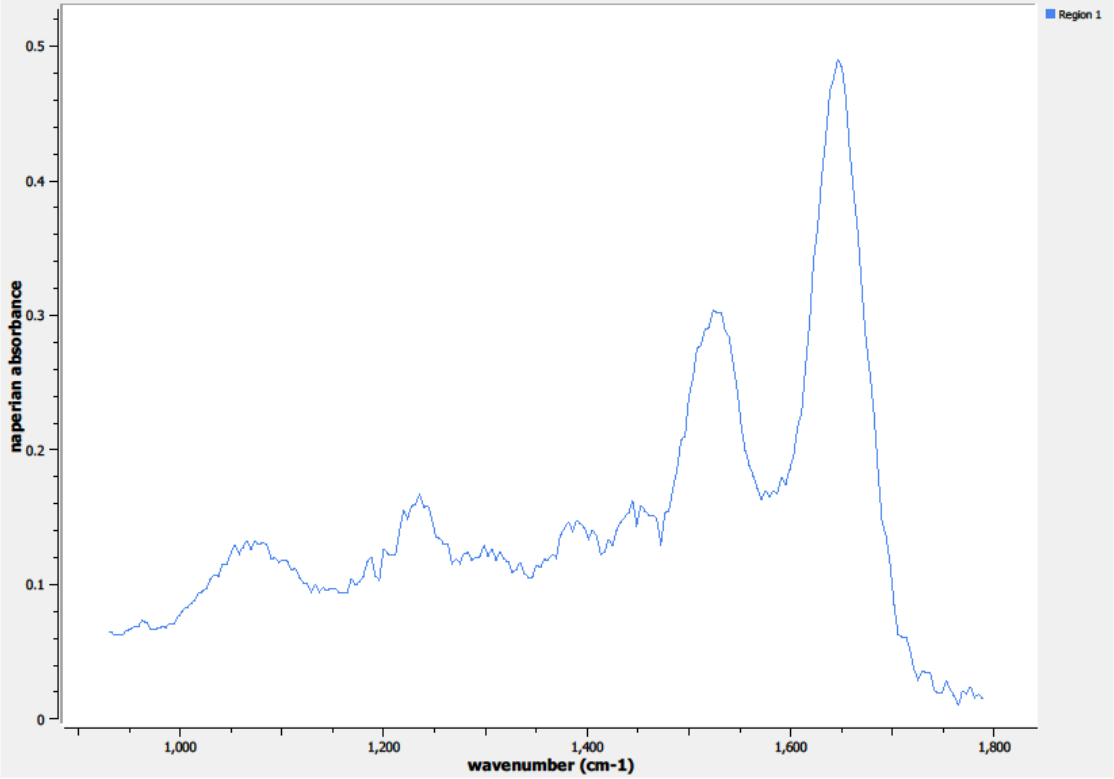



**Figure 3.64. Specimen 10-10-A177m, Point X Spectra**

## 3.4. Analysis

We used the two-tailed T-test to evaluate statistical differences in our results since deviations could be in either direction. Statistical differences were considered to be significant at $p < 0.05$. All spectra and analysis figures shown were obtained from the mean spectra of the individual specimen groups (normal or malignant,

ductal or lobular, etc.) and each hormonal and protein receptor groups (estrogen, progesterone, and HER2). Spectra and analysis values shown represent the entire experiment.

### 3.4.1. Intensity Ratio Parameter Analysis

Absorbance peak values in the "fingerprint" region similar to the cervical specimens were used for intensity ratio parameter analysis on the breast samples. The peaks were chosen by a combination of observation and taking into consideration which peaks were evaluated from the literary FTIR review described in Chapter 1. All ratios utilized the Amide II peak (1546 $cm^{-1}$) as the denominator and the other selected peak values as the numerator.

The observed peak that is the most dominant in absorbance is the protein Amide I peak at 1658 $cm^{-1}$, which is mainly attributed to C=O stretching vibration. The protein Amide II peak is also significant, which arises from C-N stretching vibration coupled with N-H bending vibration.

Both bands at 1450 $cm^{-1}$ and 1402 $cm^{-1}$ arise from CH3 bending modes of methyl proteins and lipids; 1450 $cm^{-1}$ is attributed to asymmetrical bending, and 1402 $cm^{-1}$ is attributed to symmetrical bending. The band at 1334 $cm^{-1}$ is much less prominent that others, but it does show slight increased absorption. This band is in the peptide group and is evident of normal mode symmetry and C-H vibration.

Amide III is another common protein peak used for investigations, which is around 1234 $cm^{-1}$. This absorption band is mainly due to asymmetrical phosphate

($PO_2^-$) stretching modes of nucleic acids and vibrations of collagen. The absorption peak at 1082 $cm^{-1}$ is also impacted by nucleic acids and vibration modes in collagen; in this band, though, the impact is due to symmetrical stretching of $PO_2^-$. The weak absorption seen at 1030 $cm^{-1}$ is attributed mainly to the C-O vibrational mode and the C-O stretching/bending of glycogen.

The final absorbance band we evaluated was a weak absorbance seen at 962 $cm^{-1}$. This band is in the methyl group, and the absorbance is mainly caused by $PO_4$ stretching. We decided to include this band in our analysis because we noticed a very slight absorbance seen for the malignant specimens that did not appear to be present in the normal specimens.

#### 3.4.1.1. Normal Comparison: Ethnicity

The first comparison which we investigated was evaluating white and black breast specimens to see if we could successfully differentiate between ethnicity groups. No other diagnostic histological information was provided by CHTN with the normal breast specimens. Slight variances between white and black ethnicity are depicted in the infrared spectra seen in Figure 3.65, with greater absorbance throughout the spectra of the black breast specimens.

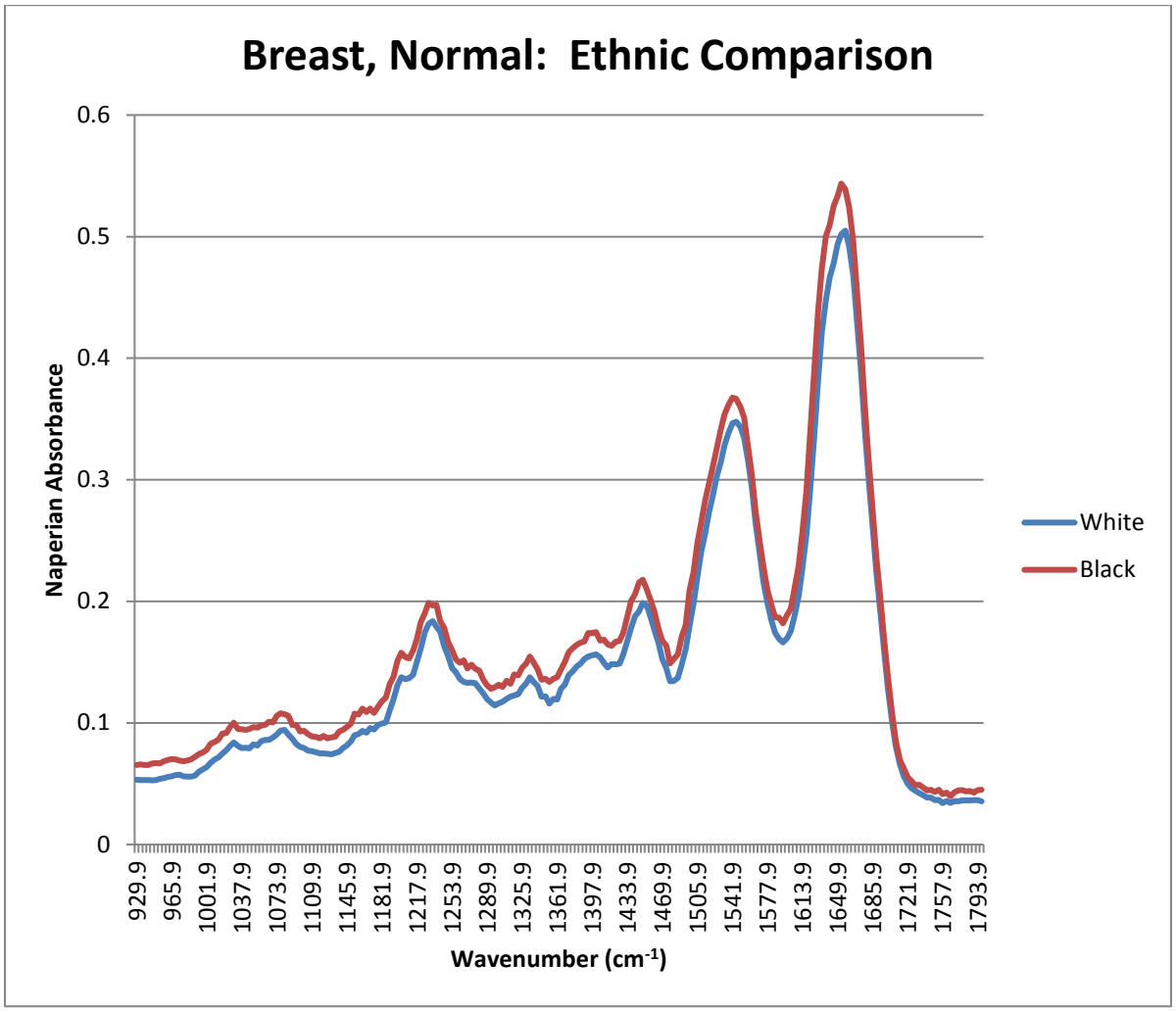



**Figure 3.65. IR spectra of ethnic groups, average of normal breast specimens**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 3.65 are shown in Figure 3.66. The graph shows the differences at peak intensity values of various bands in the spectra specific to the two ethnic groups being assessed. Although no visible absorbance peak was observed at 962 $cm^{-1}$, we included this wavenumber in our analysis because a weak absorbance peak was observed in the malignant specimens.

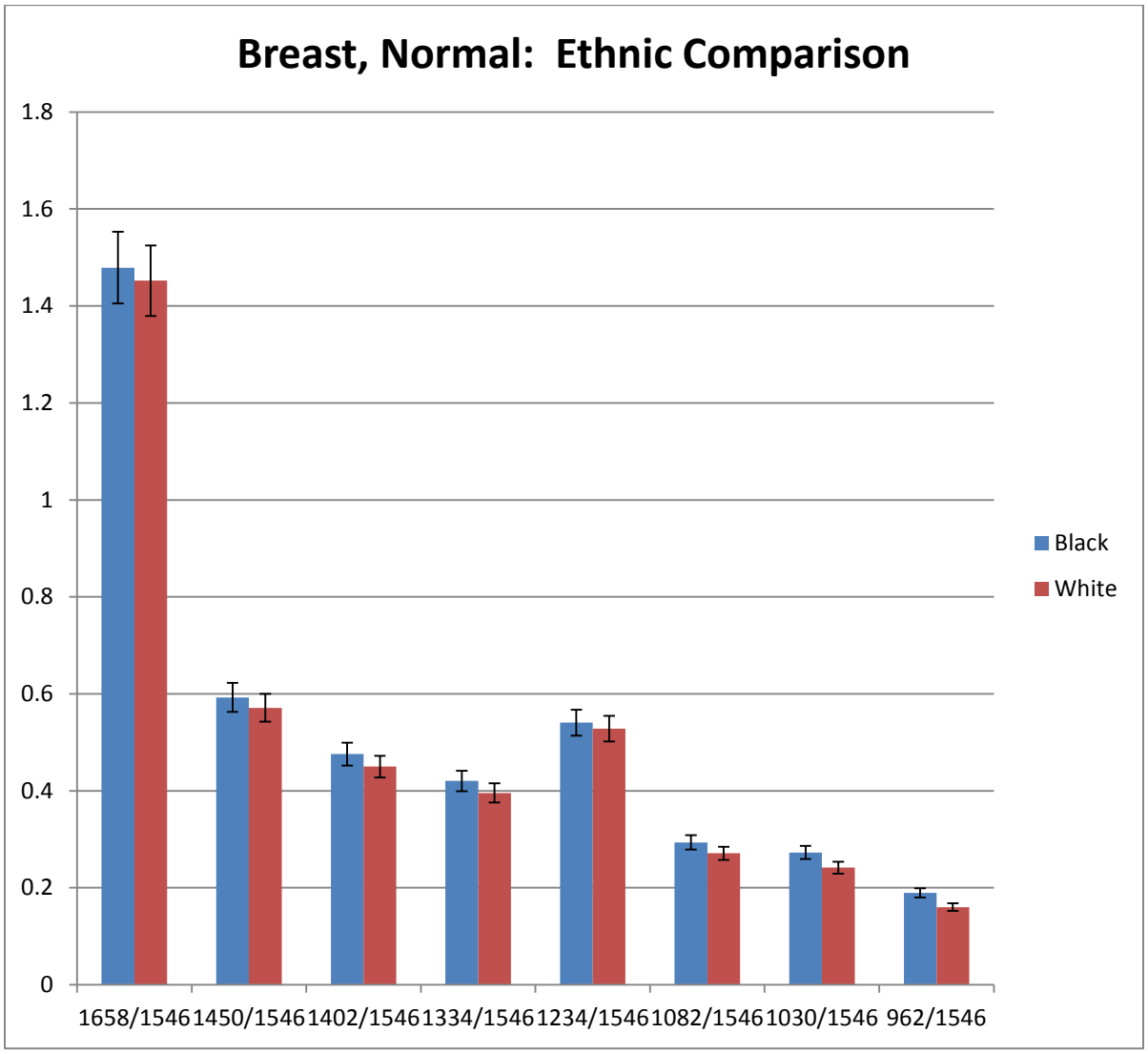



**Figure 3.66. Mean intensity ratio parameter of normal breast specimens at prominent absorbance bands throughout the fingerprint region; ethnic comparisons between black and white tissue groups**

Only two intensity ratio groupings showed statistically significant results differentiating between white and black normal breast specimens, and only one of these ratios was at an absorbance peak. One of the collagen and glycogen bands, 1030/1546, showed the only significant dissimilarities between the two ethnic

groups. In our data collection, although this lower wavenumber is not a dominant absorbance peak, it does display the greatest and only significant spectral differentiation between black and white specimens in intensity ratio parameter analysis, with the intensity ratio decreasing with the white tissue specimens. Although not an absorbance peak, the methyl band at 962/1546 also displayed significant difference between the ratios of the two ethnic specimen groupings.

We also performed intensity ratio parameter analysis with 1654 as the ratio denominator, which yielded similar results. The only significant ratios were still 1030/1654 and 962/1654, although the ratio differential was not as substantial as the 1546 $cm^{-1}$ denominator set.

The intensity ratios showed lower absorption for the white specimens for all the seven ratios which we analyzed, which is visually consistent with the spectra seen in Figure 3.65. We expect to be able to differentiate between our preliminary black and white specimens by utilizing the area under the curve analysis.

#### 3.4.1.2. Malignant Comparison: Ductal, Lobular, and Phylloides

We also evaluated the ductal, lobular, and phylloides tissue samples on normal breast specimens to see if we could successfully differentiate. Variances are portrayed in the mid-infrared spectra seen in Figure 3.67. We noted that any comparison was preliminary in this analysis that involved phylloides specimens for the reason that we had an insufficient quantity of phylloides samples available for

our spectral data collection; we only received one specimen from CHTN that was categorized as phylloides.

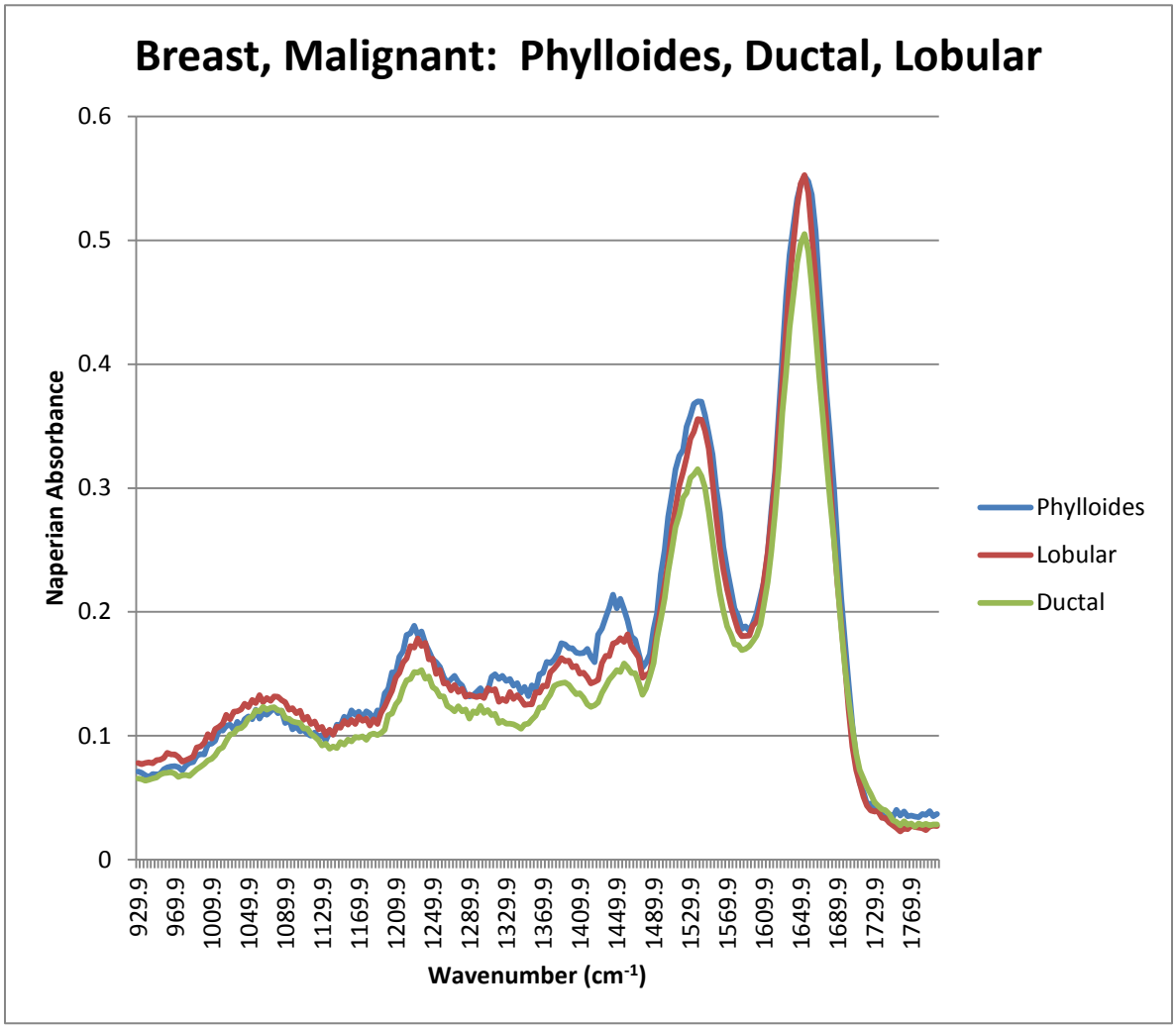



**Figure 3.67. IR spectra of average of malignant breast specimens, comparing phylloides, lobular, and ductal tissue samples**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 3.67 are shown in Figure 3.68. The plot shows the differences at peak

intensity values of various bands in the spectra specific to the three tissue types being assessed.

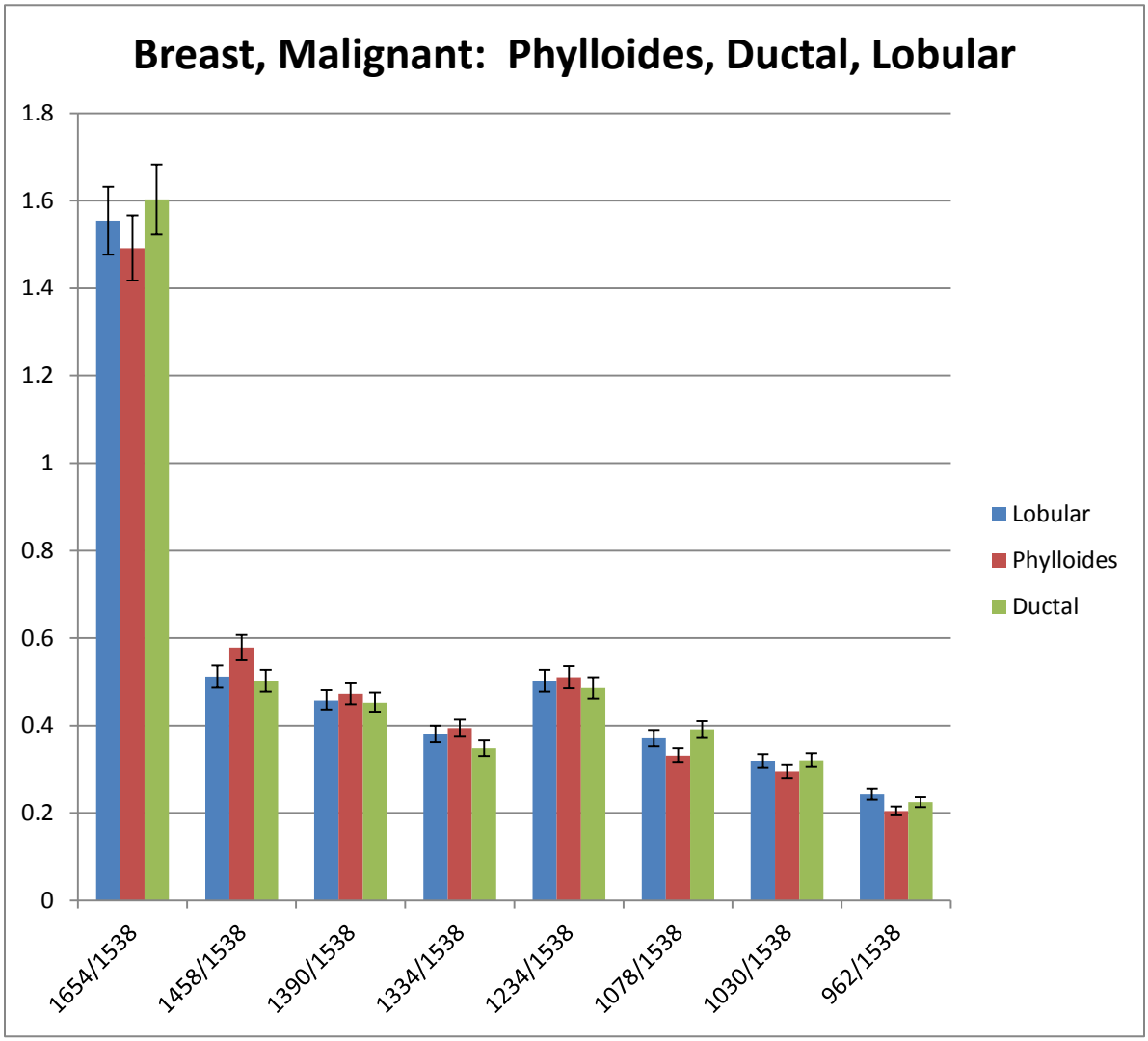



**Figure 3.68. Mean intensity ratio parameter of malignant breast specimens at different bands throughout the fingerprint region, comparing lobular, ductal, and phylloides tissue with Amide II as the ratio denominator parameter**

Although most of three tissue groupings have similar ratios at the eight absorbance wavenumbers studied, and no ratio exists that can be utilized to differentiate all three tissue groupings, some statistically significant differences do exist in four of the ratios that can be used to preliminarily glean discerning factors among the three tissue groupings. One of the more prevalent ratios is due to methyl proteins and lipids at 1458/1538, which shows differences between the phylloide specimen with both the ductal and lobular specimens. However, the delta at this ratio is not significant enough to distinguish between the ductal and lobular tissues.

The weak absorbance seen in the peptide band does display some significant ratio differences at 1334/1538. The distinguishable ratios are only between phylloide and ductal specimens, however; the lobular tissue samples are not distinguishable between either the ductal or phylloide specimens.

The nucleic acid and collagen band at 1078 $cm^{-1}$ also shows similar results as the methyl protein and lipids ratio; the 1078/1538 ratio differentiates between the phylloide specimens and the ductal and lobular specimens. Again, this ratio is unable to distinguish between lobular and ductal specimens.

Finally, methyl band at 962 $cm^{-1}$ reveals some significant results. Although the ductal specimens are indistinguishable between either lobular or phylloides tissues, the ratio does distinguish between lobular and phylloide. While this difference can be identified in the 1458/1538 ratio, this ratio provides another supporting data point.

In the investigation of comparing phylloides, ductal, and lobular malignant specimens, the Amide I band as the denominator in the ratio analysis actually presents slightly better results than the Amide II band in the denominator value, which can be seen in Figure 3.69. Five of the ratios present some statistical significance compared to the four ratios utilizing the Amide II denominator previously discussed.

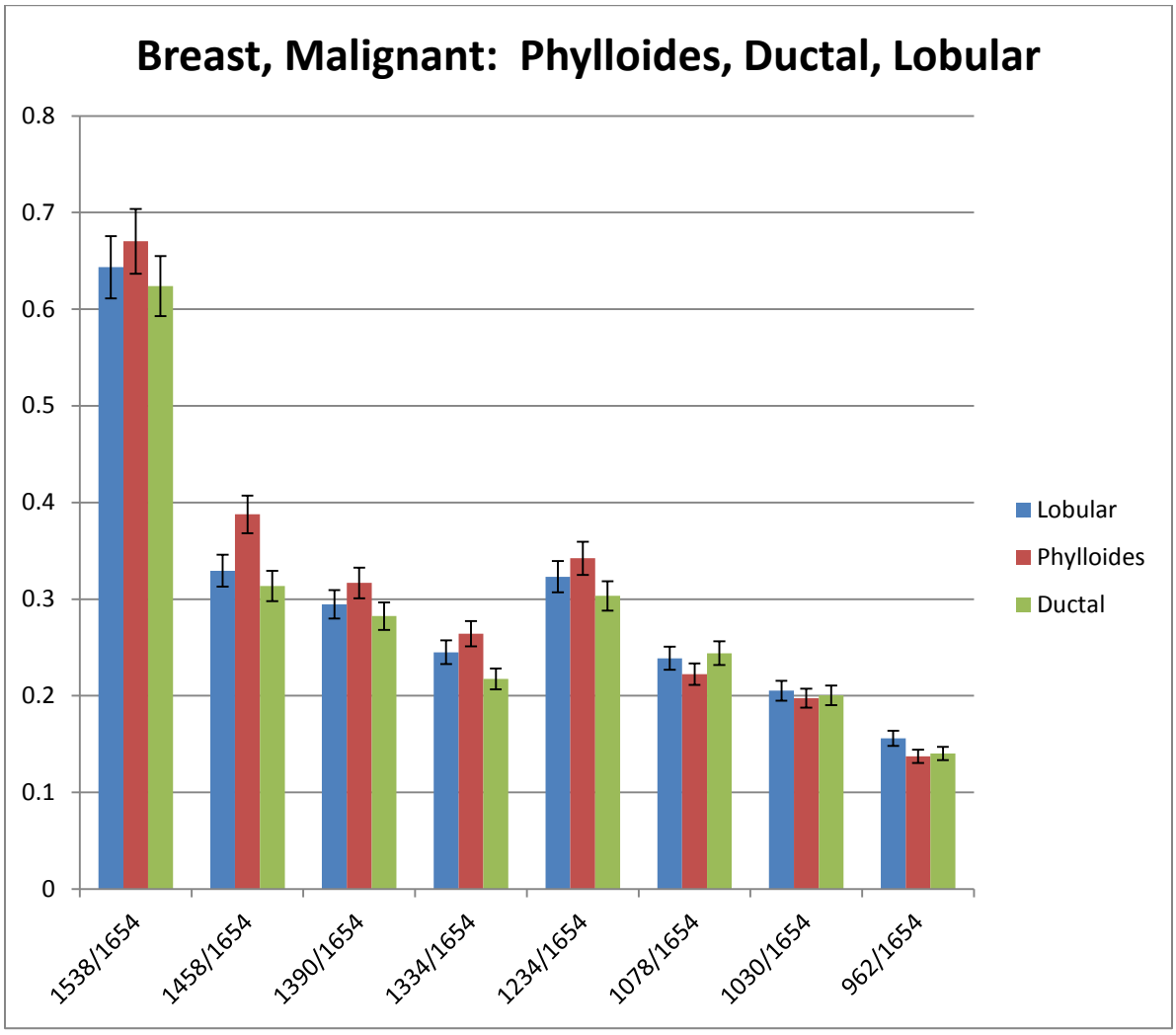



**Figure 3.69. Mean intensity ratio parameter of malignant breast specimens at different bands throughout the fingerprint region, comparing lobular, ductal, and phylloides tissue with Amide I as the ratio denominator parameter**

The methyl protein and lipids bands again display the greatest prevalence in this analysis; utilizing the Amide I band as the denominator, however, makes both 1458/1654 and 1390/1654 ratios significant. The 1458/1654 ratio again displays differences between the phylloide and the ductal specimens as well as phylloide and

lobular specimens; however, statistical significance is not demonstrated between the ductal and lobular tissues. The other ratio that utilizes the methyl protein and lipid band at 1390/1654 distinguishes between ductal and phylloide tissue. This ratio is not able to differentiate lobular tissue from the other two tissue types, though.

The weak absorbance seen in the peptide band also displays significant ratio differences at 1334/1654. The ratios distinguish between ductal and phylloide specimens and ductal and lobular specimens. 1334/1654 is the only ratio that allowed ductal tissue samples to be discernable from both lobular and phylloide samples.

The ratio with the nucleic acid and the collagen Amide III protein band as the numerator parameter also showed results of some significance. This absorbance intensity ratio at 1234/1654 differentiates only between the ductal and phylloide specimens; lobular tissue samples do not show significant ratio differences that distinguish from the other two tissue groupings.

The final ratio that showed any significance was with the methyl band 962 $cm^{-1}$ in the numerator. This ratio showed significant differences between the lobular and both the phylloide and ductal specimens, which provides greater distinction than with the 1538 $cm^{-1}$ denominator data set.

### 3.4.1.3. Malignant Comparison: Receptors

After comparing the malignant breast tissue types (ductal, lobular, and phylloide), we evaluated the hormonal and protein differences present in these specimens. The three primary markers utilized in breast cancer diagnostics are the hormonal receptors estrogen (ER) and progesterone (PR) and the protein receptor HER2; the associated specimens' spectra are seen in Figure 3.70. Being limited by the diagnostics provided by CHTN with our specimens, we only evaluated the following three receptor categories: ER+/PR+/HER2-, ER+/PR-/HER2-, and ER-/PR-/HER2- (triple negative).

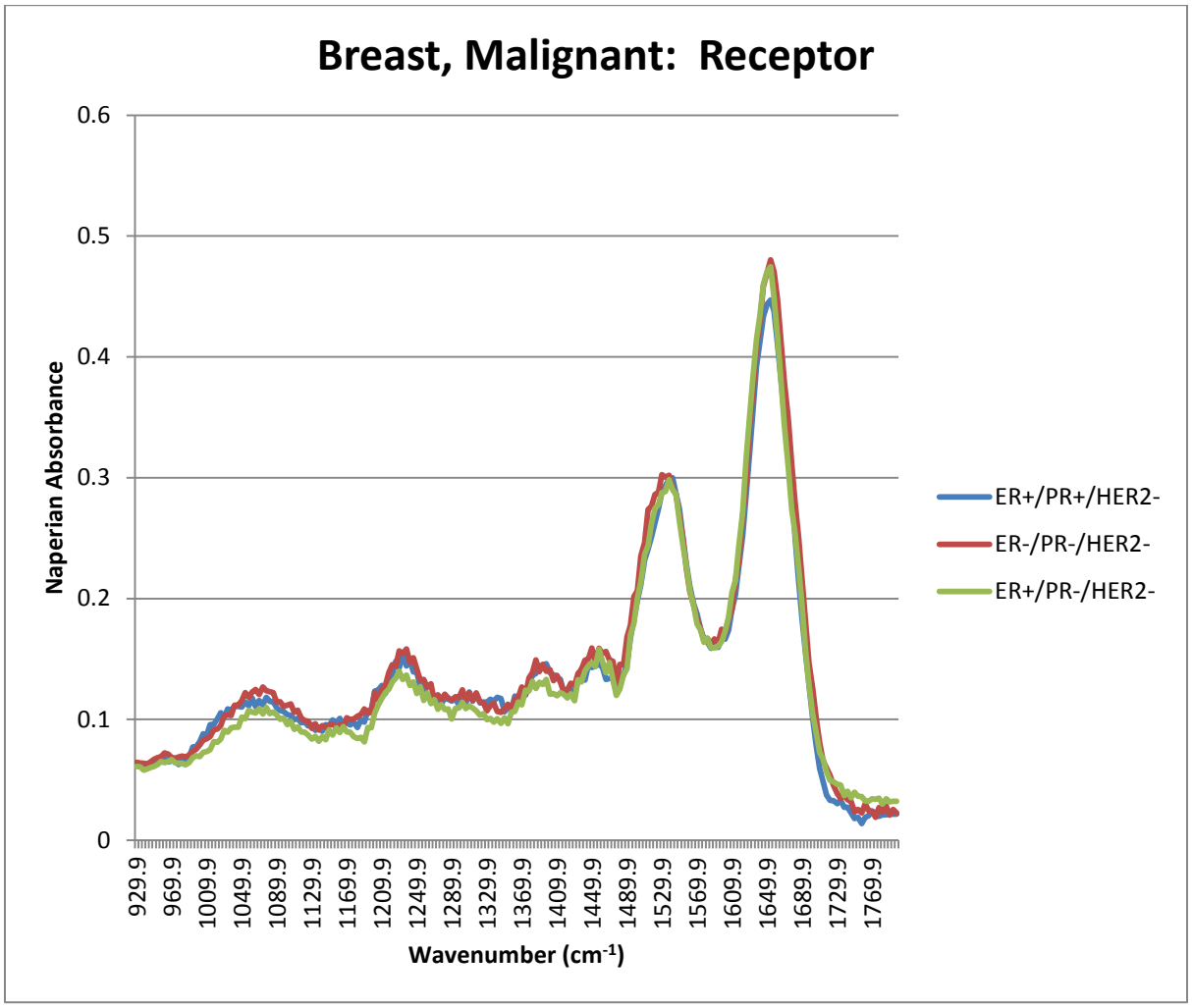



**Figure 3.70. IR spectra of average of malignant breast specimens, comparing hormonal and protein receptors**

A visual comparison of the infrared spectra of the three receptors revealed no remarkable observable differences in the data. The triple negative specimen (ER-/PR-/HER2-) appears to have slightly higher absorbance in the collagen and Amide III bands, although the analysis will reveal definitive variances.

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 3.70 are shown in Figure 3.71. The plot shows the differences at peak intensity values of various bands in the spectra specific to the three hormonal and protein receptors being assessed. As with our previous breast analysis, we initially utilized the Amide II peak at 1538 $cm^{-1}$ in the ratios' denominator, then performed another intensity ratio parameter analysis with the Amide I peak at 1654 $cm^{-1}$ in the denominator.

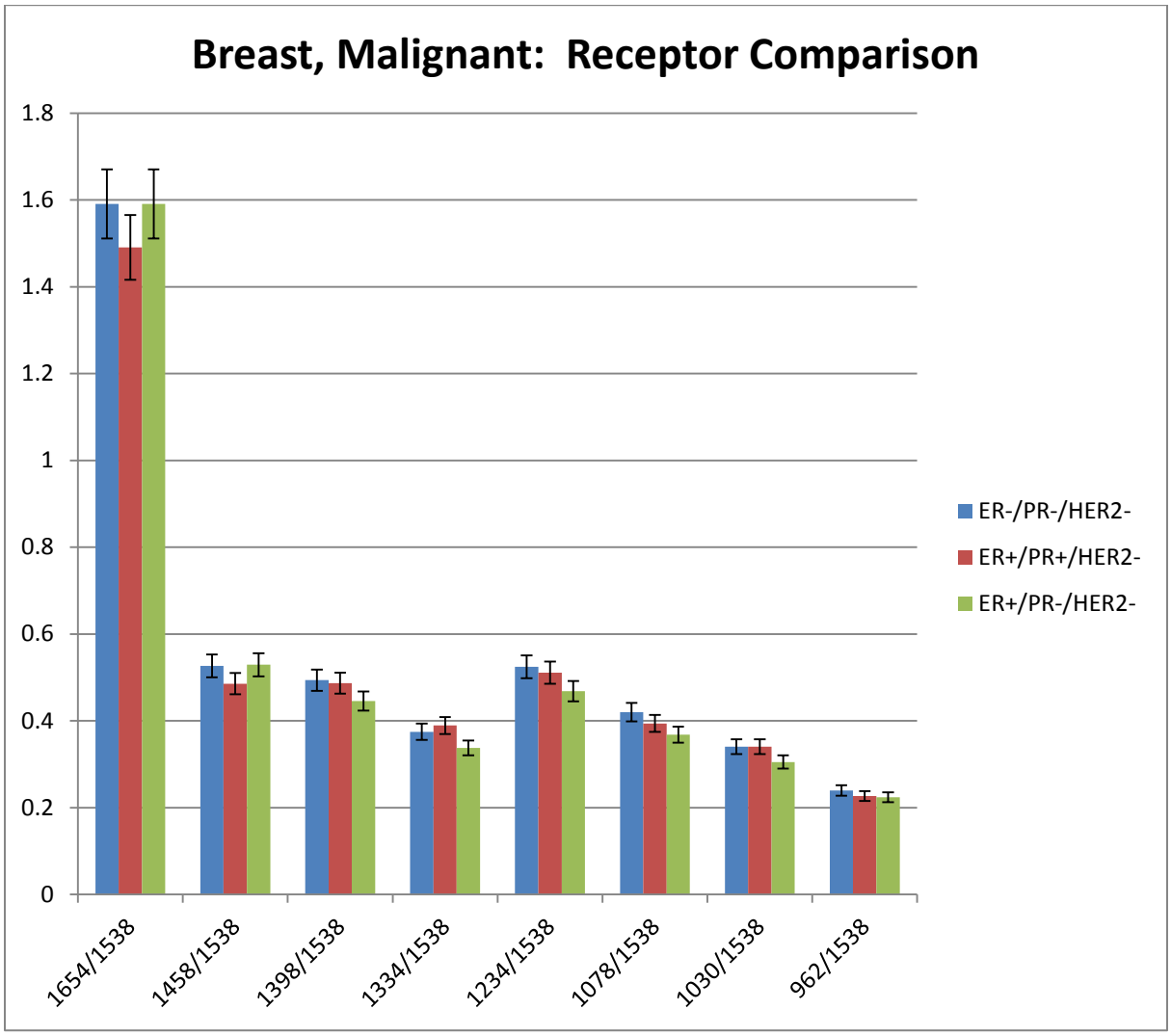



**Figure 3.71. Mean intensity ratio parameter of malignant breast specimens with Amide II as denominator at different bands throughout the fingerprint region, comparing specimen diagnosis of the two hormonal receptors estrogen (ER) and progesterone (PR) and the protein receptor HER2**

Although most of three tissue groupings have similar ratios at the eight absorbance wavenumbers studied, and no ratio exists that can be utilized to differentiate all three tissue groupings, some statistically significant differences do exist in five of the ratios that can be used to glean discerning factors among the

three tissue groupings. The two collagen ratios show the greatest distinguishing ratios between specimen groupings.

The absorbance seen in the methyl protein and lipid band and the Amide III protein band display ratios at 1398/1538 and 1234/1538, respectively, that distinguish between the triple negative specimen and the specimens diagnosed as ER+/PR-/HER2-. In both cases, the triple negative tissue sample cannot be separated from the ER+/PR+/HER2- specimens with significance.

The weak absorbance ratio at 1334/1538 due to the peptide group shows significant differentiation between the tissues diagnosed as ER+/PR-/HER2- with both the triple negative specimen and the specimens diagnosed as ER+/PR+/HER2-. We were not able to successfully detect between the triple negative and ER+/PR+/HER2- specimens at this intensity ratio parameter.

The two collagen groups at 1078 $cm^{-1}$ and 1030 $cm^{-1}$ showed greater differentiation between the ratios. We were able to discriminate between the triple negative specimen and the ER+/PR-/HER2- specimens by utilizing the 1078/1538 ratio, although we could not distinguish the ER+/PR+/HER2- samples from either of the other two. The weak collagen absorbance at the 1030/1538 ratio showed difference from the ER+/PR-/HER2- tissue samples with both the triple negative and ER+/PR+/HER2- specimens. As with the 1398/1538 ratio, though, we could not also differentiate between the triple negative and ER+/PR+/HER2- specimens.

In the peak intensity ratio analysis of receptor comparison, the Amide I band as the denominator in the ratio analysis actually presents substantially better results than the Amide II band in the denominator value, which can be seen in Figure 3.72. Although we found statistically significant results in only five of the eight ratios, similar to the Amide II denominator analysis, the results were much improved.

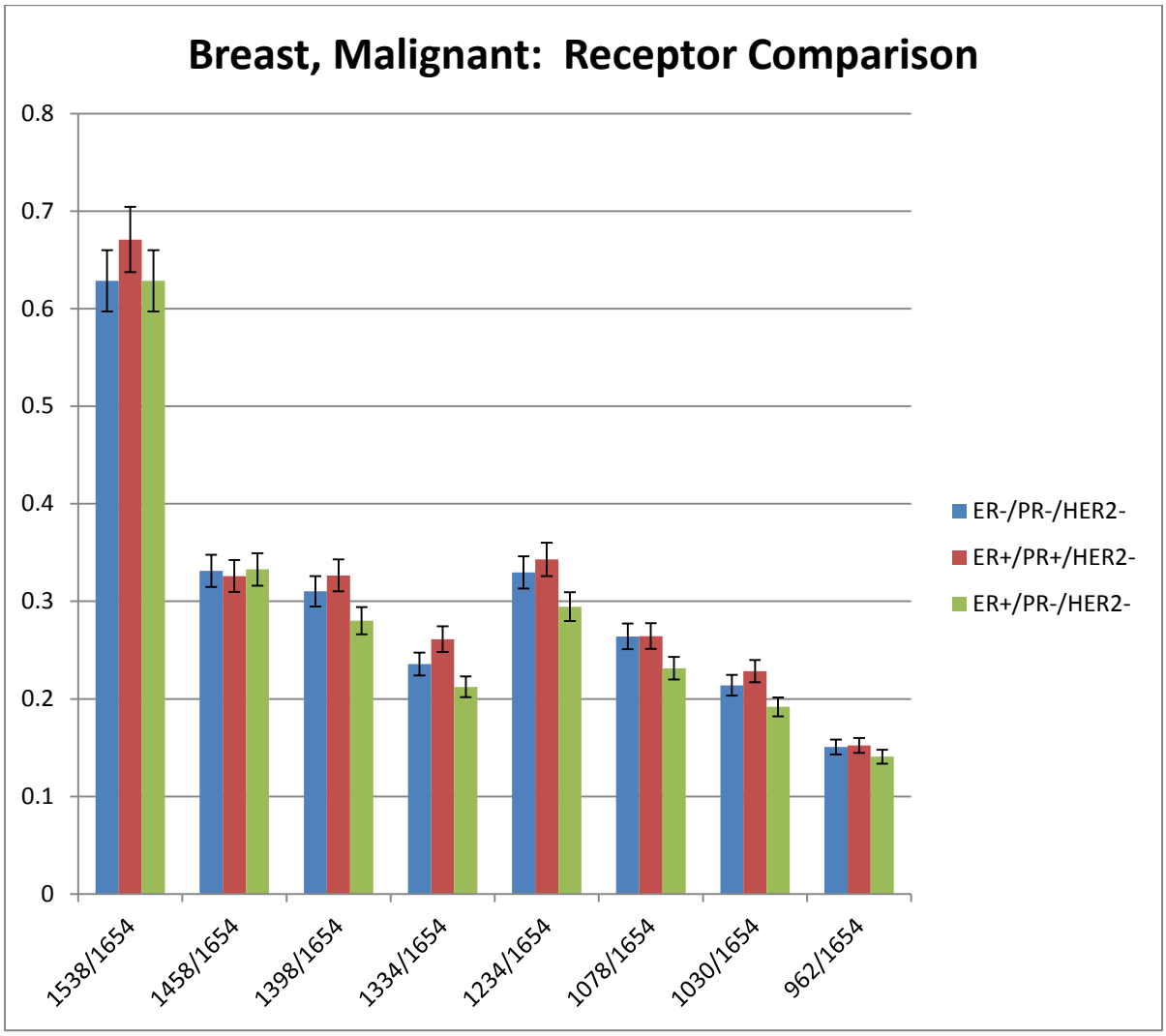



**Figure 3.72. Mean intensity ratio parameter of malignant breast specimens with Amide I as denominator at different bands throughout the fingerprint region, comparing specimen diagnosis of the two hormonal receptors estrogen (ER) and progesterone (PR) and the protein receptor HER2**

We were able to gather relevant information from the ratio analysis in five of the eight ratios that we used to extract discerning factors among the three types of specimens being analyzed. The same ratios that utilized Amide II in the denominator contained relevant information; however, greater differences existed

and more of the tissue samples were discernable within these ratios. In fact, for the first time we were actually able to separate the triple negative specimen from the other two tissue specimen types.

We were not able to utilize either the Amide II protein band at 1538 $cm^{-1}$ or the methyl protein and lipid band with asymmetrical bending at 1458 $cm^{-1}$ ratio to differentiate specimens with different receptor diagnosis. However, the methyl protein and lipid band with symmetrical bending at 1398 $cm^{-1}$ did utilize relevant information. We were able to identify ER+/PR-/HER2- samples from the triple negative and ER+/PR+/HER2- specimens; we were not, however, able to differentiate the triple negative and ER+/PR+/HER2- specimens from each other.

The Amide III protein band ratio at 1234/1654 also showed significant differences between ratios to allow us to differentiate between ER+/PR-/HER2- and the triple negative and ER+/PR+/HER2- specimens but not between the triple negative and ER+/PR+/HER2- specimens. Likewise, both collagen band ratios at 1078/1654 and 1030/1654 revealed the same results of being able to identify the samples diagnosed as ER+/PR-/HER2- from the other two specimen types but not being able to differentiate between the triple negative specimen and the tissue samples diagnosed as ER+/PR+/HER2-.

The final ratio showing significant differences is 1334/1654, which is the weak absorbance due to the peptide group. Not only is this ratio the only one that we were able to utilize to segregate the triple negative specimen, we were also able

to segregate the other two specimen diagnosis types. We were still able to clearly differentiate ER+/PR+/HER2-, ER+/PR-/HER2-, and ER-/PR-/HER2- from each other with significance.

We were anticipating the possibility of being able to identify specific influences from each hormone and protein receptor. However, based on the results we gathered and developed in the intensity ratio parameter analysis, we were unable to extract data specific to the individual receptors.

#### 3.4.1.4. Malignant Comparison: Ethnicity

We next compared the mean spectral data between malignant specimens ethnically grouped by CHTN. Similar to the normal specimens, the expectation of the data sets had been very similar spectra between the two malignant sample types, but overall absorbance intensity can be observed in Figure 3.73 between the white and black infrared spectra. We anticipate this observed variation to be most prevalent in the area-under-the-curve analysis (refer to Section 3.4.2), although we still expect relevant results in the intensity ratio parameter analysis; greater absorbance is observed throughout the spectra of the white specimens.

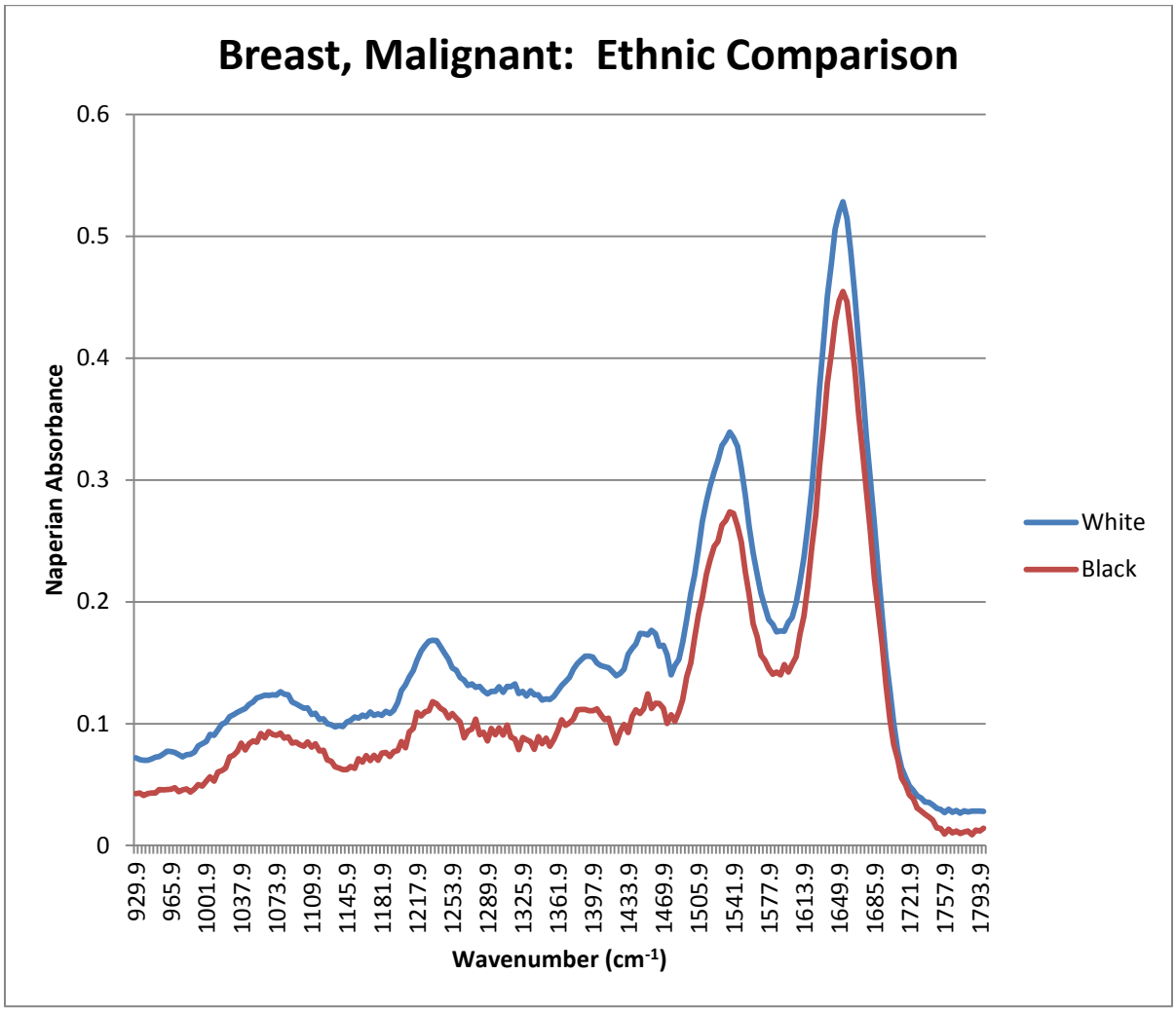



**Figure 3.73. IR spectra of average of malignant breast specimens, comparing ethnicity**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 3.73 are shown in Figure 3.74. The plot shows the differences at peak intensity values of various bands in the infrared spectra specific to the two ethnic groups being assessed. Several of the ratios displayed in the plot show significant

differences between white and black malignant specimens, which, in comparison, only two ratios in the normal comparison displayed similar substantial differences.

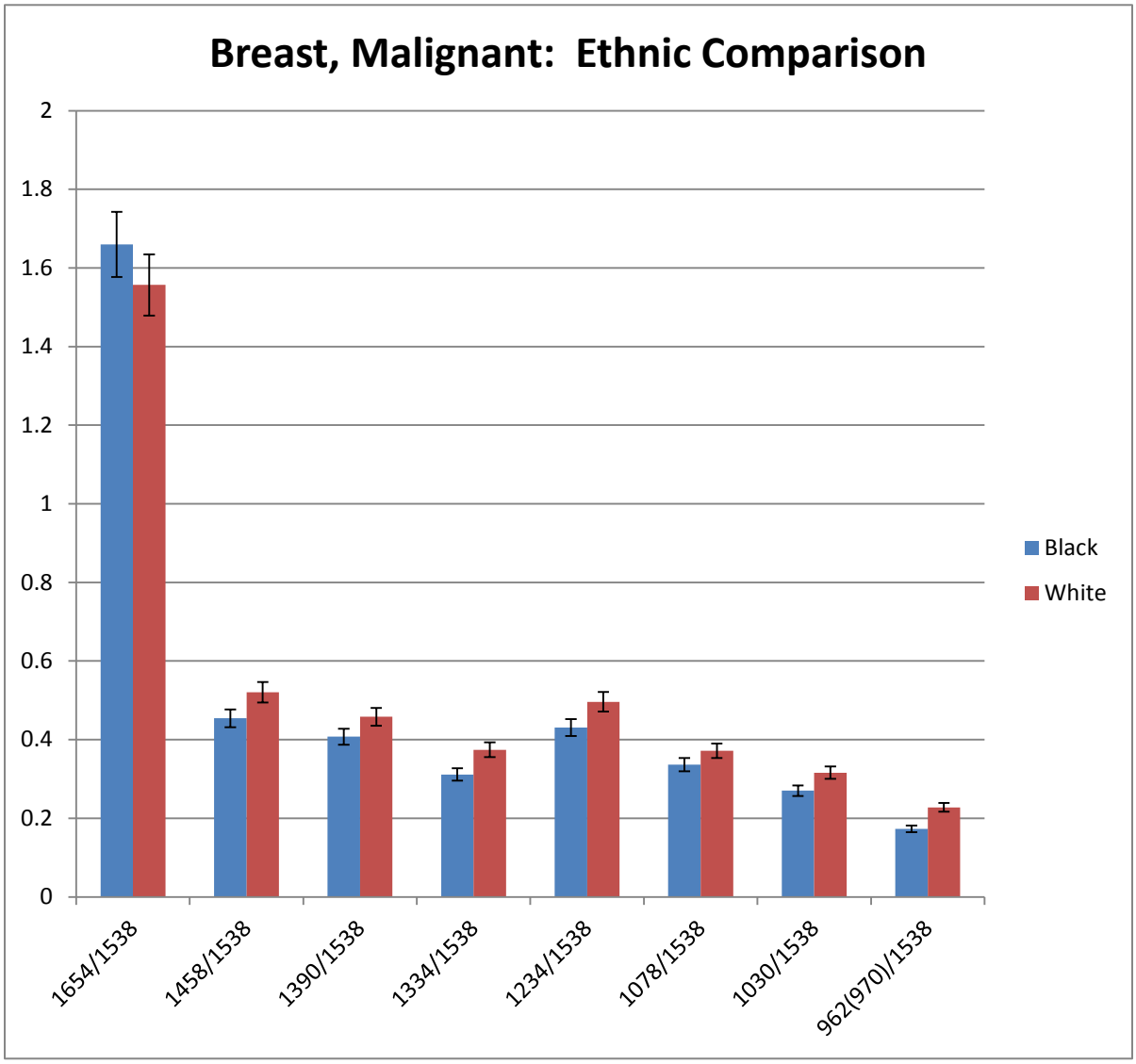



**Figure 3.74. Mean intensity ratio parameter of malignant breast specimens at prominent absorbance bands throughout the fingerprint region; ethnic comparisons between black and white tissue specimens**

Seven of the eight bands analyzed revealed significant ratio differences. The only group that did not show significant difference was the 1654/1538 band, which was due to the Amide I/Amide II protein bands. All the other bands demonstrated an ability to significantly discriminate between white and black malignant breast specimens.

Although we performed intensity ratio parameter analysis with both the Amide I and Amide II protein bands as the denominator, and each analysis data set showed the same set of ratios as showing differences significant enough to be used to differentiate white and black malignant specimens, the differences between white and black specimen ratios were much greater with Amide I as the denominator, as can be seen in shown in Figure 3.75. Similar to the Amide II denominator analysis set, the only band that did not show significance between white and black specimen mean intensity absorbance ratios was 1538/1654, which was due to the Amide II/Amide I protein bands. All the other bands demonstrated an ability to differentiate with significance between white and black malignant breast specimens.

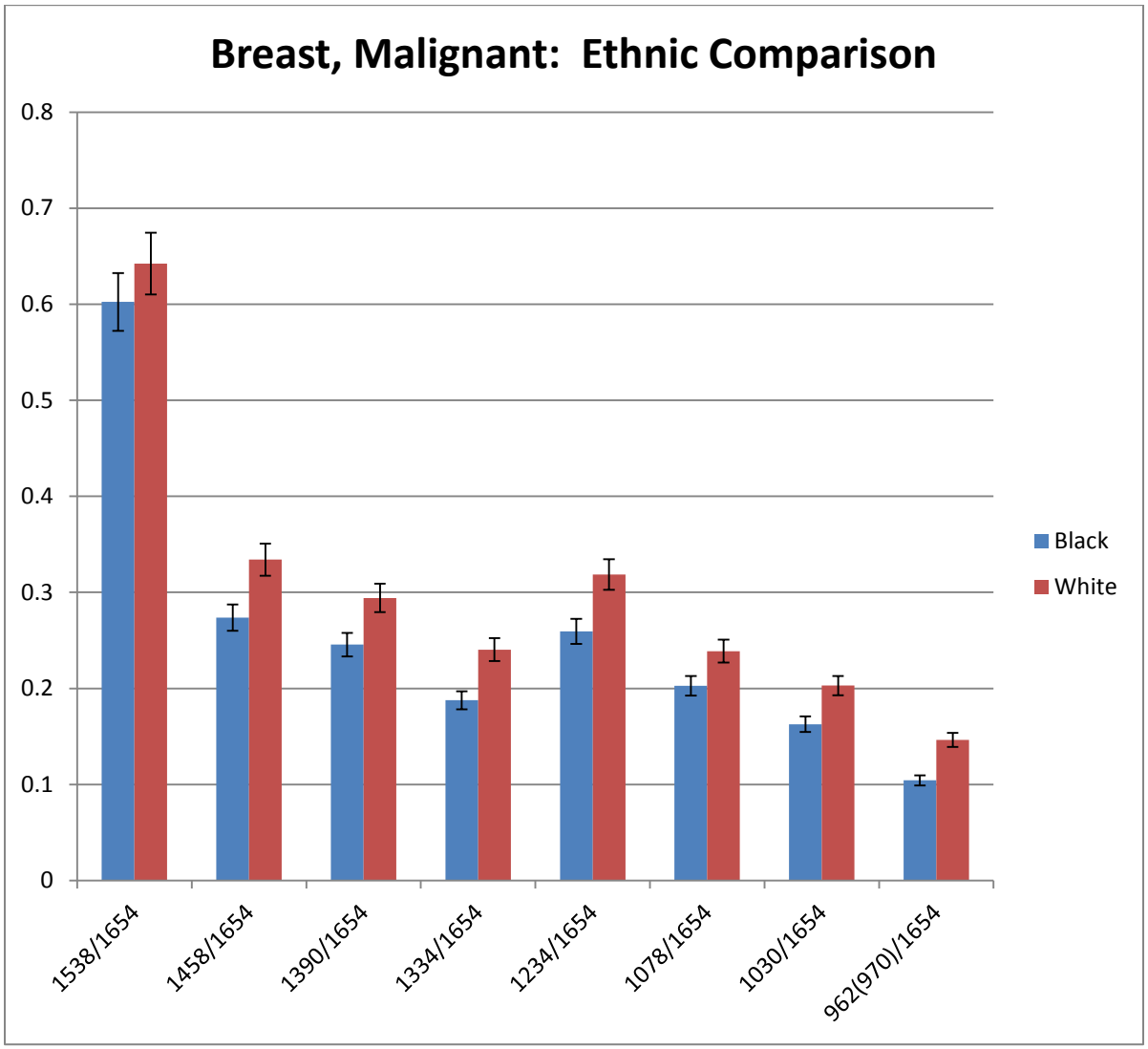



**Figure 3.75. Mean intensity ratio parameter of malignant breast specimens at prominent absorbance bands throughout the fingerprint region; ethnic comparisons between black and white tissue specimens**

The spectra showed considerably higher absorbance for the white malignant specimens, and the mean ratios also showed consistently higher value for each white specimen ratio analyzed. We also expect to be able to clearly differentiate

between these two ethnic groups by utilizing the area under the curve analysis in Section 3.4.2.

#### 3.4.1.5. Breast Tissue Comparison: Normal vs. Malignant

The final intensity ratio comparison we investigated was evaluating normal versus malignant breast specimens. During this analysis, we compared the ethnicity of normal and malignant; we also specifically looked at the triple negative specimen due to the diagnostic difficulties associated with these receptor readings. In the normal and malignant comparison, we evaluated the following collections: normal white vs. malignant white, normal black vs. malignant black, and cumulative normal vs. cumulative malignant. We then extracted the triple negative results from the cumulative malignant breast data and compared that specimen data to the rest of the malignant data and the cumulative normal data.

##### 3.4.1.5.1. Ethnicity

After analyzing breast specimen ethnicity within normal tissue samples and malignant tissue specimens, we compared the ethnicity between normal and malignant samples. The comparisons were broken down as depicted in Figure 3.76.

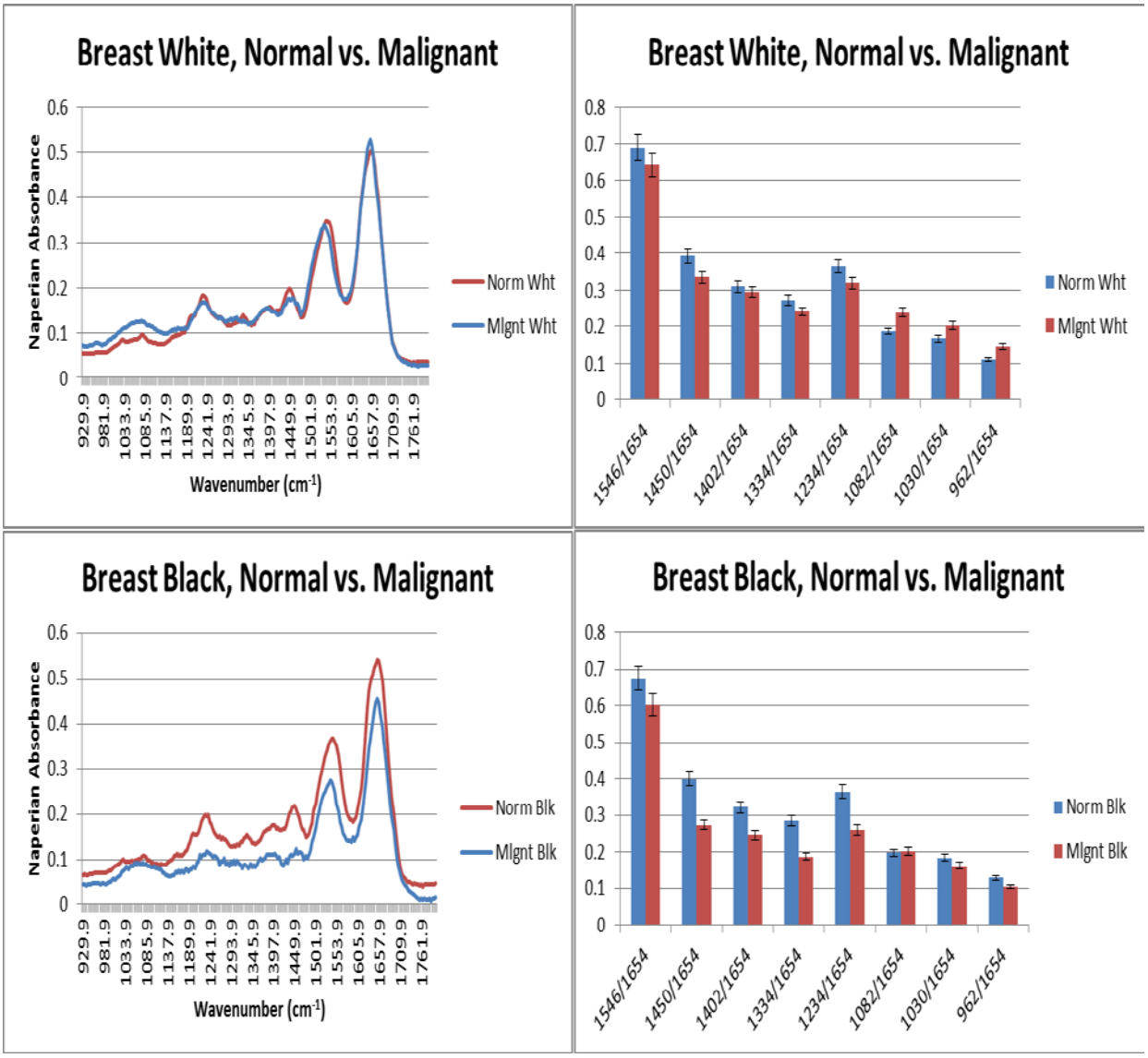



**Figure 3.76. Normal and malignant breast tissue comparison. (Left Column) IR spectra of normal and malignant breast specimens. (Right Column) Mean intensity ratio parameter of normal and malignant breast specimens at different bands throughout the fingerprint region.**
**ROWS: (Top) White breast specimens. (Bottom) Black breast specimens.**

Comparing the infrared spectra of the white breast specimens reveals similar spectra throughout the evaluated range with the exception of the collagen and

glycogen bands. Essentially, in the range below the Amide III protein band, the white normal samples have lower absorbance than the white malignant samples. The normal black samples, however, have significantly higher absorbance throughout the entire spectra compared to the malignant black samples. We expect the results to be more evident in the Area-Under-the-Curve analysis performed in Section 3.4.2.

The intensity ratio parameter analysis shows six ratios as having significant differences between malignant and normal white specimens when utilizing the Amide I protein band as the denominator. Two ratios affected by methyl proteins, 1450/1654 and 960/1654, allow us to differentiate between normal and malignant white samples. We were not able to distinguish between normal and cancer samples utilizing the symmetrical bending of the methyl proteins and lipids at 1402/1654, however.

We were also able to identify malignant from normal white tissue for both collagen and glycogen ratios at 1082/1654 and 1030/1654. In addition, the peptide group at 1334 $cm^{-1}$ and the Amide III protein at 1234 $cm^{-1}$ also allow us to identify the malignant samples from the normal samples. The only other identifying ratio we were not able to utilize was the Amide II protein at 1546/1654. These results were expected, though, as this ratio has not proved to be as useful in identification throughout our intensity ratio parameter analysis.

When we analyzed these ratios with the Amide II protein band in the denominator, the results only revealed the two collagen and glycogen bands at 1082 $cm^{-1}$ and 1030 $cm^{-1}$ and the one methyl protein band at 962 $cm^{-1}$ to show any significant differences between normal and malignant tissue. The other five ratios did not demonstrate statistically significant differences.

Evaluating the black tissue specimens revealed much more distinction between the normal and malignant tissue groupings when employing the Amide I protein band in the denominator. In fact, we were able to use all but one of the ratios to distinguish between normal and malignant samples. The only ratio that did not have a significant difference was the nucleic acid and collagen band at 1080/1654. All other seven ratios evaluated in this intensity ratio parameter analysis showed significant distinction between the normal and malignant black tissue samples, and the differences between ratios showed much greater variances between normal and malignant than we saw with the white specimens. When we analyzed these ratios with the Amide II protein band in the denominator, the results revealed two ratios that could not be used to distinguish, the collagen and glycogen band at 1030 $cm^{-1}$ and the methyl protein band at 962 $cm^{-1}$.

#### 3.4.1.5.2. Cumulative

After comparing the normal and malignant white samples and the normal and malignant black samples, we analyzed all the normal and malignant breast samples cumulatively, the spectra of which can be seen in Figure 3.77. Our

expectations were that we would clearly be able to distinguish between normal and malignant specimens.

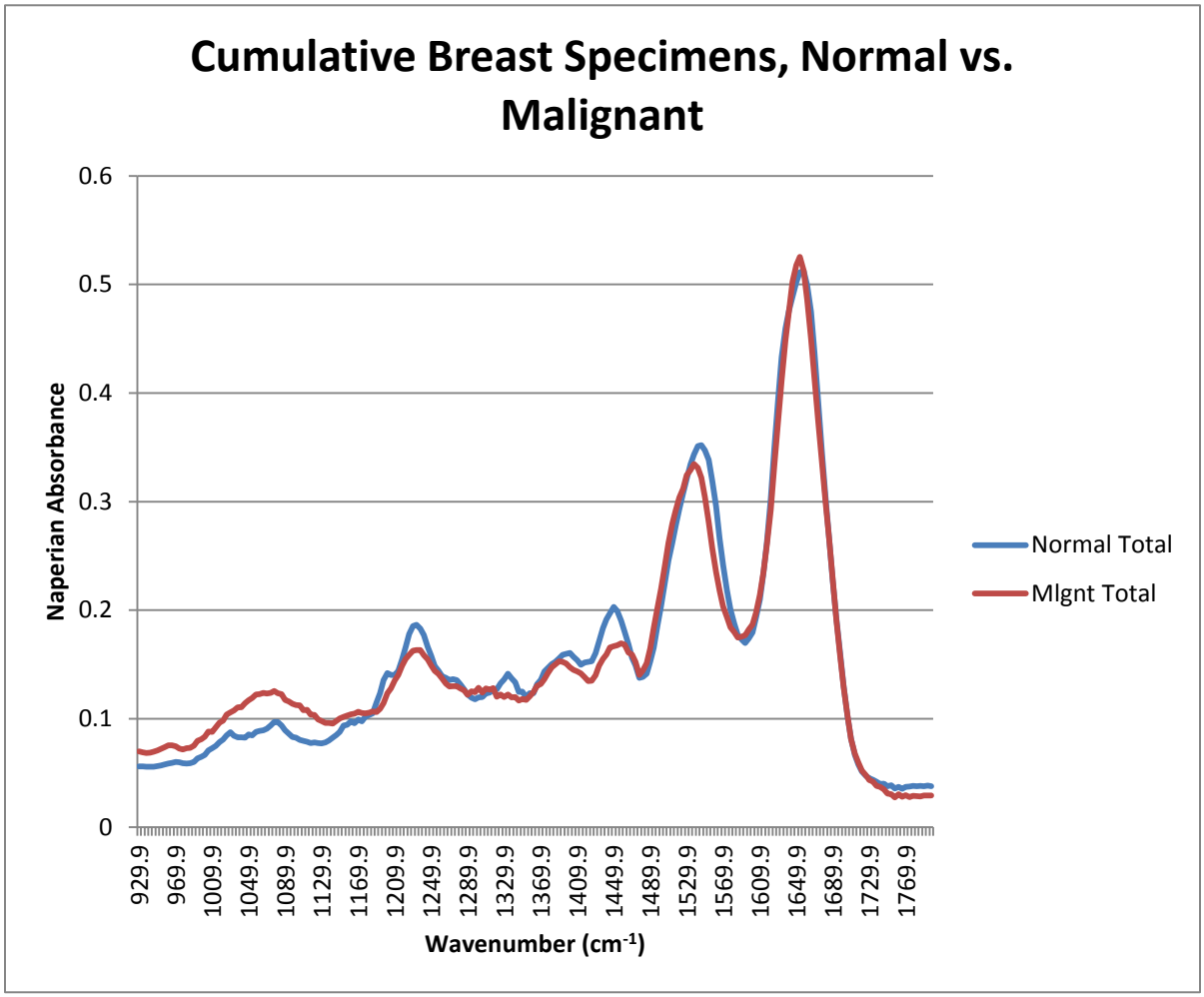



**Figure 3.77. IR spectra comparing total cumulative normal and malignant breast specimens**

The ratios of the mean intensity absorbance peaks seen in the spectra in Figure 3.77 are shown in Figure 3.78. The plot shows the differences at peak intensity values of various bands in the infrared spectra specific to the normal and

malignant specimens being assessed. The collagen, glycogen, and methyl protein bands in the lower wavenumbers below Amide III reveal higher absorbance for the malignant samples; however, the higher wavenumbers from Amide III through Amide I show higher absorbance for the normal specimens. We expect the results to be more evident in the Area-Under-the-Curve analysis performed in Section 3.4.2.

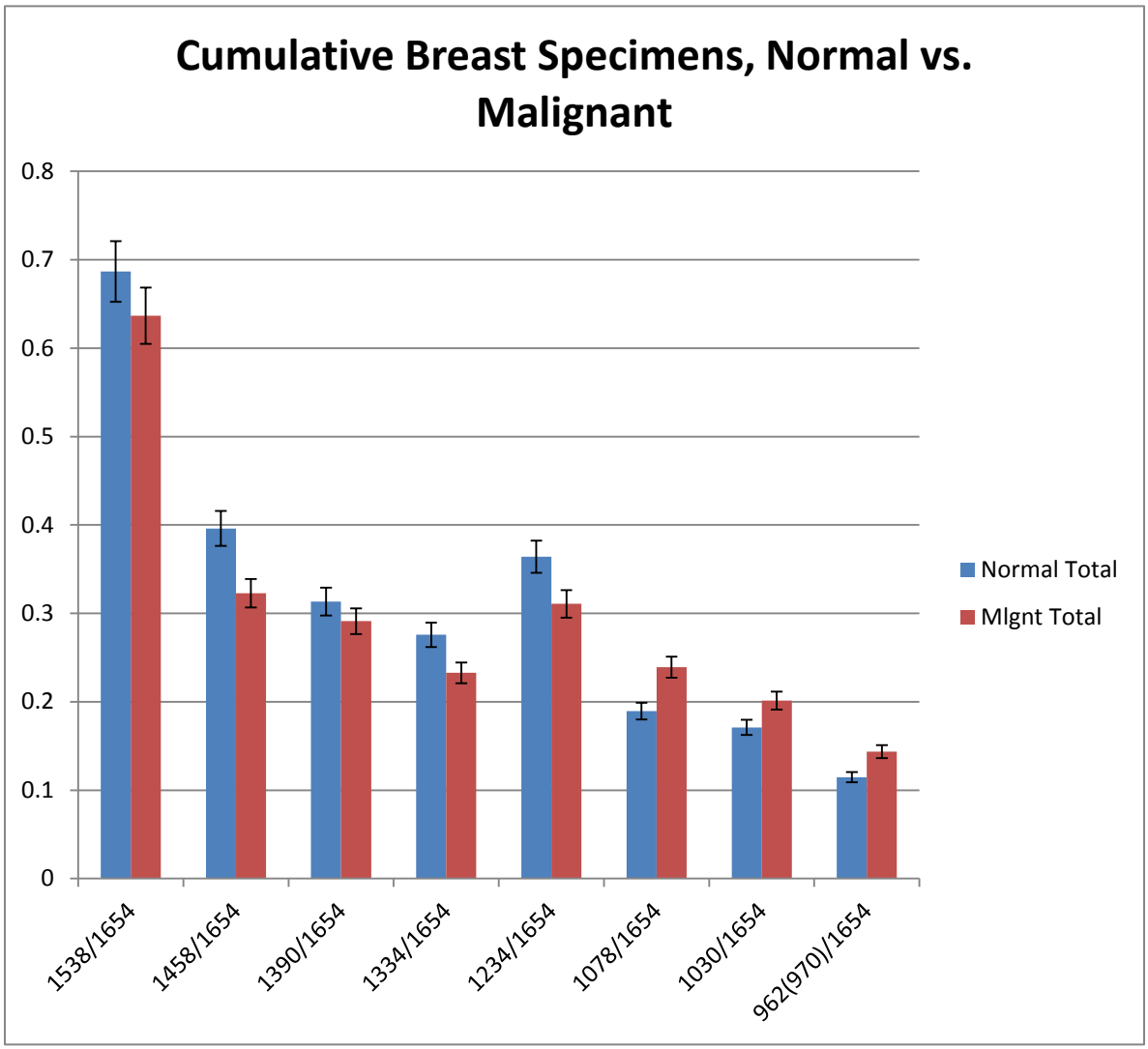



**Figure 3.78. Mean intensity ratio parameter comparing normal and malignant breast specimens at prominent absorbance bands throughout the fingerprint region**

When utilizing the Amide I peak as the ratio's denominator, six of the eight bands analyzed revealed significant ratio differences. The only two ratios that did not show significant differences between normal and malignant samples were the Amide II protein band at 1538 $cm^{-1}$ and the symmetrical bending methyl protein at 1390 $cm^{-1}$.

Two ratios affected by methyl proteins, 1458/1654 and 962/1654, allowed us to differentiate between normal and malignant specimens. In fact, the band at 1458 $cm^{-1}$ showed the greatest significance in clear differentiation of ratios of absorbance peaks. As expected from the absorbance spectra seen in Figure 3.77, the methyl protein band at 962 $cm^{-1}$ showed the malignant ratio to be higher than the normal ratio.

We were also able to clearly identify malignant from normal tissue for both collagen and glycogen ratios at 1078/1654 and 1030/1654, respectively, which again, consistent with the IR spectra, revealed higher peak intensity ratios for the malignant samples. The final two peak intensity ratio parameters we utilized to differentiate between normal and malignant tissue samples were the peptide group at 1334 $cm^{-1}$ and the Amide III protein at 1234 $cm^{-1}$.

All six of the peak intensity ratio parameters previously identified with Amide I in the denominator had clear and consistent significant differences between normal and malignant specimens. When we analyzed these ratios with the Amide II protein band in the denominator, however, the preliminary results were not as

distinguishable as with Amide I as the denominator. Only four of the eight parameters demonstrated any statistically significant differences between normal and malignant tissue, which were the two methyl protein parameters at 1458/1538 and 962/1538 and the two collagen and glycogen influences at 1078/1538 and 1030/1538.

#### 3.4.1.5.3. Triple Negative

The final intensity ratio parameter analysis comparison we made focused specifically on the one triple negative tissue specimen due to the diagnostic challenges associated with tissue with these receptor indicators. Specifically, comparing the triple negative specimen with the normal breast specimens was the focus to evaluate the diagnostic capabilities of IR spectroscopy; we also compared the results to all the non-triple negative malignant specimens. In this analysis of the triple negative specimen, all references to "malignant specimens" or "malignant samples" are referring to all malignant specimens except the triple negative sample, unless otherwise specified. Again, since we were only able to acquire one triple negative specimen from CHTN, our results at this point are very preliminary.

Slight visual variances are depicted in the absorbance intensity seen in Figure 3.79, which are the mean spectra of the malignant and normal tissue specimens and the triple negative spectrum. The average absorbance can be seen to be greater in the normal tissue throughout the fingerprint region with the exception of the collagen and glycogen bands around 1025 $cm^{-1}$ and 1080 $cm^{-1}$, which is

significantly lower than the malignant samples. We anticipated these bands would show the greatest significance when comparing the intensity ratio parameters.

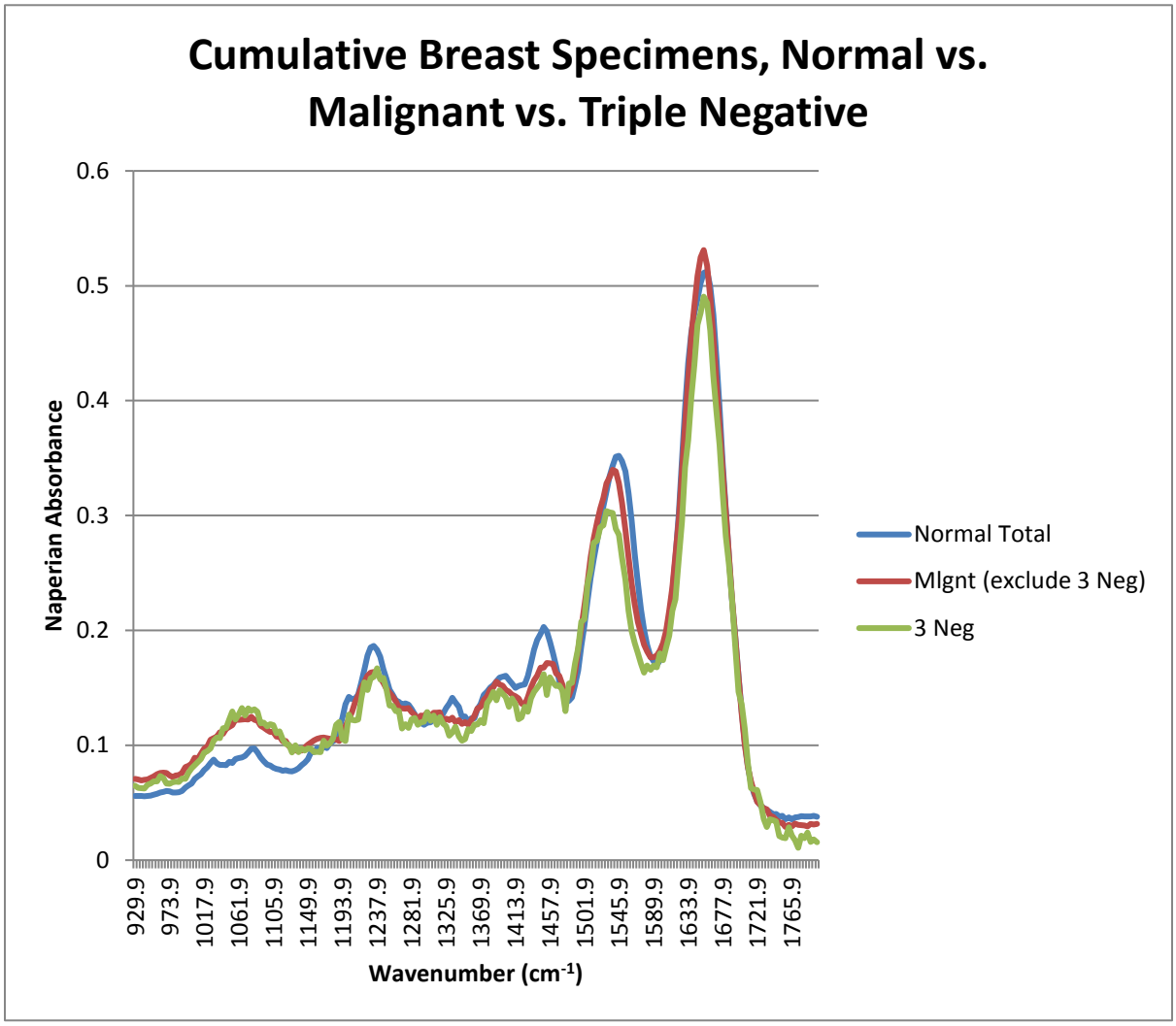



**Figure 3.79. IR spectra comparing average absorbance of normal, malignant (excluding triple negative), and triple negative breast tissue specimens**

The ratios of the primary mean intensity absorbance peaks seen in the spectra in Figure 3.79 are shown in Figure 3.80. We first utilized Amide II as the

denominator value. The plot shows the differences at peak intensity ratio values of various bands in the spectra specific to the normal and malignant tissue samples. Three of the eight peak ratios distinguish between the normal specimens and the triple negative specimen; one shows significant difference among all three groups (normal, malignant excluding triple negative, and triple negative).

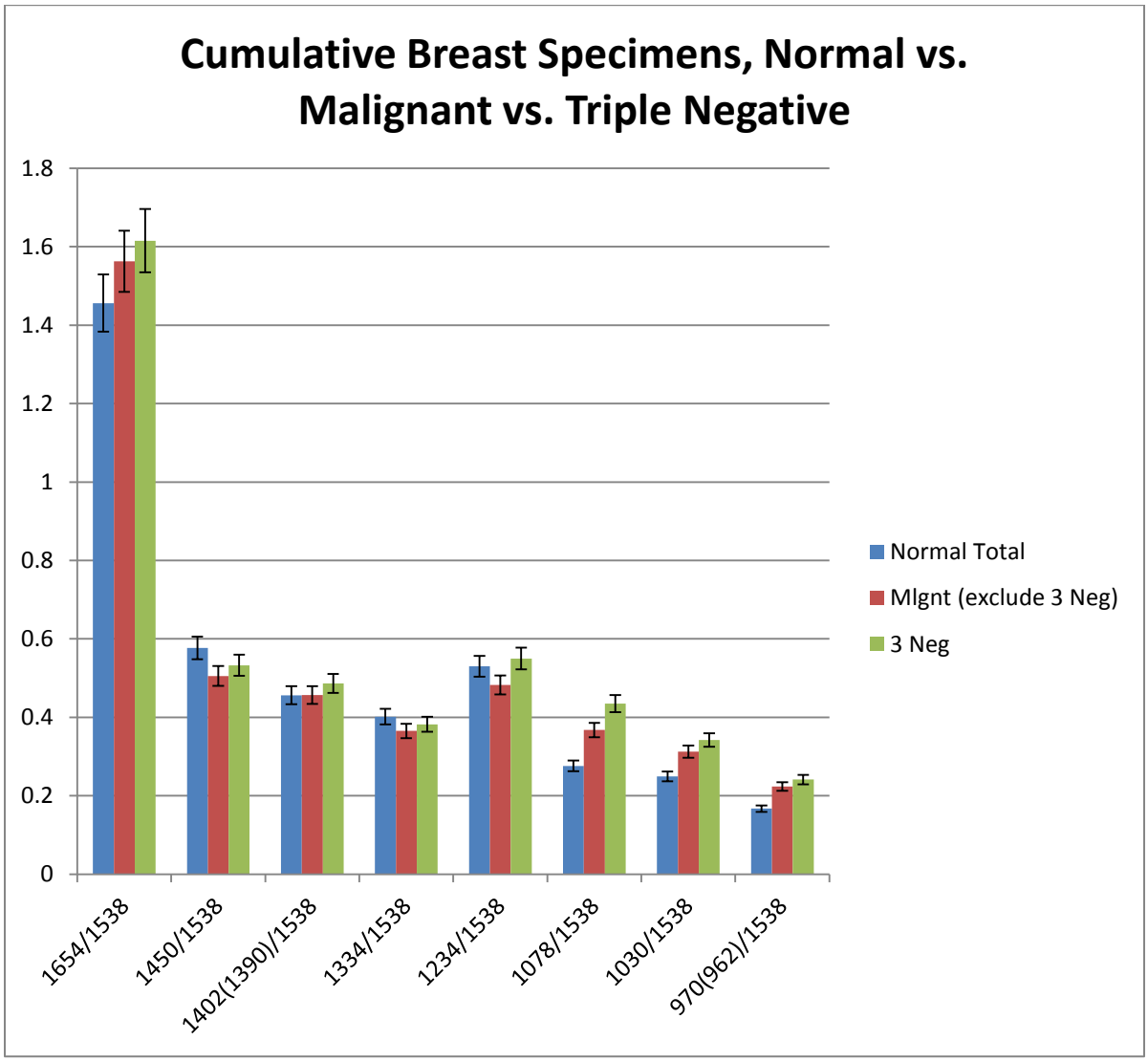



**Figure 3.80. Mean intensity ratio parameter of breast specimens at different bands throughout the fingerprint region for normal, malignant (excluding triple negative), and triple negative specimens with the Amide II protein as the denominator parameter**

The asymmetrical bending methyl protein and lipid ratio, 1450/1538, differentiated between the normal and malignant specimens. The Amide III protein ratio at 1234/1538 also showed significant difference between the malignant and

triple negative specimens. No other significant differences were observed in these two ratios.

Both the methyl protein ratio at 970/1538 and the glycogen ratio at 1030/1538 distinguished between normal and malignant specimens and, more importantly, between the normal and the triple negative specimens. Furthermore, we were able to successfully differentiate among all three of the normal, malignant, and triple negative specimens by utilizing the collagen ratio band at 1078/1538. The results of these three ratios were expected due to the variances seen in the Figure 3.79 spectra.

After performing intensity ratio parameter analysis with the Amide II protein as the denominator parameter, we performed the same analysis with the Amide I protein in the denominator, the results which are depicted in Figure 3.81. The plot again shows the differences at peak intensity ratio values of various bands in the spectra specific to the normal, malignant, and triple negative tissue samples with differing results from the previous analysis. Five of the eight peak ratios distinguish between the normal specimens and the triple negative specimen, and one peak displays significant difference among all three groups (normal, malignant excluding triple negative, and triple negative).

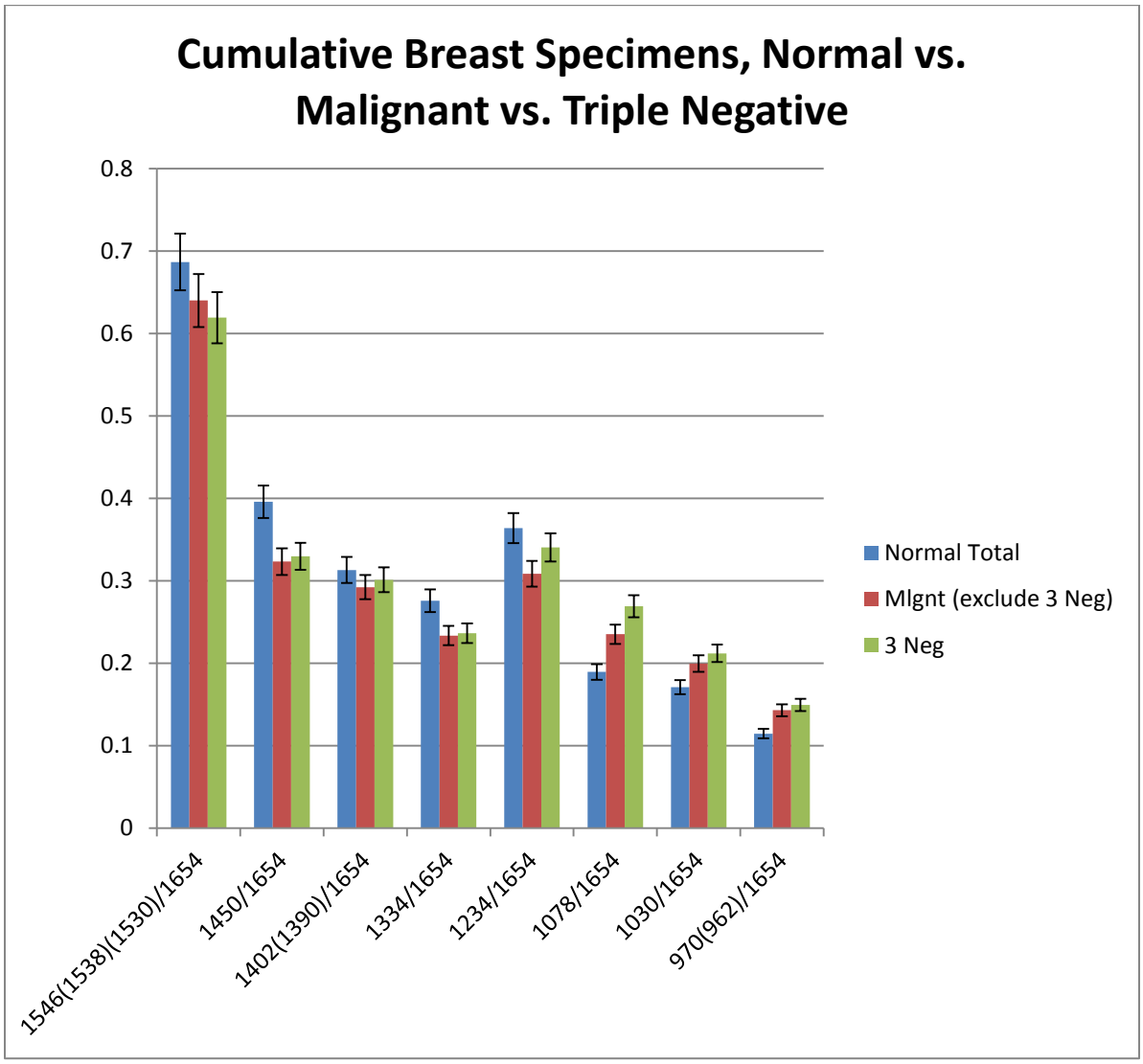



**Figure 3.81. Mean intensity ratio parameter of breast specimens at different bands throughout the fingerprint region for normal, malignant (excluding triple negative), and triple negative specimens with the Amide I protein as the denominator parameter**

The Amide III protein ratio at 1234/1654 displayed significant difference between the normal and malignant specimens and the malignant and triple negative

specimens. However, we were unable to discriminate the normal and triple negative specimens at this ratio.

We were able to use four ratios to differentiate between normal and malignant specimens and normal and triple negative specimens. The first was the 1450/1654 ratio due to asymmetrical bending of methyl protein and lipids; the second peak intensity ratio parameter was 1334/1654, which is affected by the peptide group. The third and fourth ratios were 1030/1654 and 970/1654, which are influenced by glycogen and methyl protein, respectively.

The nucleic acid and collagen band at 1078/1654 was the only ratio with the Amide I protein as the denominator factor that was able to distinguish among normal, malignant, and triple negative samples with significance. The 1078 $cm^{-1}$ absorbance band is influenced mainly by nucleic acids and collagen. The analysis results were expected from visual observation of the spectra in Figure 3.79 that portrayed differences in the spectra.

### 3.4.2. Area-Under-the-Curve Analysis

The area-under-the-curve (AUC) provides insight to absorbance intensity differences between spectra that are scaled similarly throughout the spectrum range, thus rendering the intensity ratio parameter analysis ineffective. The area-under-the-curve was calculated using the trapezoid rule.

The AUC was calculated for the spectra of all three malignant breast tissue types we investigated (ductal, lobular, and phylloides). In Table 3.3, we can see the AUC differential between the three categories.

| | Area-Under-the-Curve | | |
|---|---|---|---|
| | Ductal | Lobular | Phylloides |
| Malignant | 126.99 | 140.61 | 147.06 |

**Table 3.3. Area-under-the-curve for the mean infrared spectra of malignant ductal, lobular, and phylloides breast tissue samples.**

The malignant tissue samples show significant variance among the AUCs to be able to discriminate the ductal specimens from the lobular and phylloides specimens. However, we were unable to distinguish the lobular and phylloides samples.

The AUC was also calculated for the spectra of the three receptor categories previously analyzed, which were the two hormonal receptors estrogen and progesterone and the protein receptor HER2. Table 3.4 shows the AUC differential between the three receptors. After calculating the AUCs, we were unable to differentiate between any of the receptors. This result was not unexpected due to the similar receptor spectra graphed in Section 3.4.1.3.

| | Area-Under-the-Curve | | |
|---|---|---|---|
| | ER+/PR+/HER2- | ER+/PR-/HER2- | ER-/PR-/HER2- |
| Malignant | 119.29 | 117.45 | 124.76 |

**Table 3.4. Area-under-the-curve for the mean infrared spectra of hormonal and protein receptors in malignant breast tissue samples.**

We then calculated the AUC for normal white, normal black, malignant white, and malignant black breast tissue samples. We also included a category of all normal and all malignant breast specimens to compare. Table 3.5 displays this AUC data and the differences between these similar groupings.

| | Area-Under-the-Curve | | |
|---|---|---|---|
| | White | Black | Cumulative |
| Normal | 130.27 | 144.89 | 133.52 |
| Malignant | 135.40 | 101.29 | 133.75 |

**Table 3.5. Area-under-the-curve for the mean infrared spectra of white, black, and total number of specimens of normal and malignant breast tissue samples.**

We were able to evaluate the AUC data and isolate the normal white specimens and the normal black tissue specimens. We were also able to

discriminate the malignant black specimens from the malignant white specimens based on the AUC.

We then started assessing the normal and malignant comparisons; we were able to significantly differentiate between normal black and malignant black breast specimens. The differential between the normal and malignant black specimens was the greatest difference in our analysis. We were also able to categorize the malignant black from normal white specimens. No other identifications could be made using AUC analysis. We expected to be able to differentiate between all specimen groups based on initial visual inspection of the spectra. However, we were unable to make a distinction between normal white and malignant white specimens and, more importantly, the AUC differential between the aggregate specimens of normal and malignant breast tissues did not allow us to definitively separate.

Finally, the AUC shown in Table 3.6 was calculated for the mean infrared spectra of the all the normal and malignant breast tissue samples. The emphasis for this analysis was to distinguish the triple negative specimen from the other specimens; the AUC for the malignant samples excludes the triple negative specimen.

| | Area-Under-the-Curve | | |
|---|---|---|---|
| | Normal | Malignant (excluding 3 Neg) | Triple Negative |
| Total Breast Specimens | 133.52 | 134.98 | 126.39 |

**Table 3.6. Area-under-the-curve for the mean infrared spectra of normal and malignant breast tissue samples with emphasis on triple negative specimen.**

We were not able to identify the triple negative specimen from the normal or malignant samples using the AUC analysis technique. Although there is slightly higher absorbance throughout the IR spectra, the difference between the AUC of the triple negative specimen and the normal samples was still not considered significant.

### 3.4.3. Frequency Shift Analysis

Frequency shift analysis can only be tentatively performed with the current data set; spectral data collection with a resolution greater than 4 $cm^{-1}$ is needed to more fully exploit this analysis technique. However, some preliminary results with the current spectral data set can be defined.

No significant frequency shift occurred between any of the absorbance peaks we evaluated in the white and black normal specimens. Two malignant absorbance peaks, the methyl protein (symmetrical bending) at 1394 $cm^{-1}$ and the methyl protein band at 962 $cm^{-1}$, did display a frequency shift between the black and white samples. The malignant white specimens' absorbance at 1394 $cm^{-1}$ for the

symmetrical methyl protein blue shifted from the malignant black specimens' absorbance at 1386 $cm^{-1}$. In contrast, the absorbance at 962 $cm^{-1}$ of the methyl protein for the malignant white samples red shifted from the malignant black samples' absorbance at 970 $cm^{-1}$.

Although we were not able to identify the lobular and ductal samples from each other by using the frequency shift analysis technique, we were able to make a preliminary distinction of the phylloides samples from both the ductal and lobular samples. The phylloides absorbance at 1446 $cm^{-1}$ due to the asymmetrical bending of methyl protein and lipids is red shifted by approximately 12 to 16 wavenumbers from both the ductal and lobular samples at 1458 $cm^{-1}$ and 1462 $cm^{-1}$, respectively. The phylloides Amide III absorbance peak at 1230 $cm^{-1}$ is also red shifted from the ductal samples that have an Amide III absorbance peak at 1238 $cm^{-1}$; however, no significant frequency shift occurs between the lobular and phylloides samples. Finally, the phylloides methyl protein absorbance peak at 970 $cm^{-1}$ is blue shifted from the lobular samples that have a methyl protein absorbance peak of 962 $cm^{-1}$; however, no significant frequency shift occurs between the ductal and phylloides samples.

We were also able to preliminarily identify significant frequency shifts of the hormonal and protein receptors. The triple negative specimen, ER-/PR-/HER2-, showed an approximate 8 and 12 wavenumber red shift at both asymmetrical 1450 $cm^{-1}$ and symmetrical 1386 $cm^{-1}$ methyl protein and lipid absorbance bands, respectively, from the ER+/PR+/HER2- (1458 $cm^{-1}$ and 1398 $cm^{-1}$) and the ER+/PR-

/HER2- (1458 $cm^{-1}$ and 1398 $cm^{-1}$) specimens. At the Amide III protein absorbance peak, the triple negative specimen blue shifted by approximately 8 $cm^{-1}$ to 1238 $cm^{-1}$ from the ER+/PR-/HER2- samples, which had peak absorbance at 1230 $cm^{-1}$. Finally, the ER+/PR-/HER2- samples also blue shifted at the methyl protein wavelength by approximately 8 $cm^{-1}$ to 970 $cm^{-1}$ from both the ER+/PR+/HER2 and triple negative specimens at 962 $cm^{-1}$.

We were able to identify three absorbance peak frequency shifts in the white samples that allowed us to classify the tissue as either malignant or normal specimens. The most prevalent peak was the Amide II absorbance peak, which displayed an 8 $cm^{-1}$ red shift from the normal to the malignant spectra (1546 $cm^{-1}$ to 1538 $cm^{-1}$). A similar red shift of 8 $cm^{-1}$ occurred at the symmetrical bending of the methyl protein and lipids (1402 $cm^{-1}$ normal spectra to 1394 $cm^{-1}$ malignant spectra), whereas the asymmetrical bending band blue shifted from 1450 $cm^{-1}$ to 1458 $cm^{-1}$ in the normal to the malignant samples.

Two absorbance peaks showed frequency shifts significant enough for us to distinguish normal and malignant black specimens. Similar to the white specimens, the methyl protein symmetrical bending absorbance peak red shifted from the normal samples at 1402 $cm^{-1}$ to the malignant samples at 1386 $cm^{-1}$. On the contrary, the methyl protein absorbance peak blue shifted from 962 $cm^{-1}$ for the normal spectra to 970 $cm^{-1}$ for the malignant spectra.

When we evaluated all the normal specimens and all the malignant specimens using frequency shift analysis, we identified three absorbance peaks that shifted significantly enough for us to make preliminary assessments based on the results. The malignant asymmetrical methyl protein and lipid absorbance peak blue shifted 8 $cm^{-1}$ to 1458 $cm^{-1}$ from the normal absorbance peak at 1450 $cm^{-1}$, whereas the malignant symmetrical methyl protein and lipid absorbance peak red shifted 12 $cm^{-1}$ to 1390 $cm^{-1}$ from the normal absorbance peak at 1402 $cm^{-1}$. The malignant samples also red shifted from the normal samples most notably due to the sharp absorbance peaks for the Amide II protein; our analysis showed the shift was 8 $cm^{-1}$, from 1546 $cm^{-1}$ to 1538 $cm^{-1}$.

The final analysis we performed was looking at the frequency shifts between the mean spectra of the normal specimens and the triple negative specimen. (No significant shift was evident between the triple negative and the remaining malignant specimens.) Three absorbance peaks showed a significant red shift from the normal spectra to the triple negative spectrum. The triple negative absorbance peaks for the symmetrical methyl protein and lipid and the nucleic acid and collagen both red shifted from the normal spectra by 8 $cm^{-1}$, from 1402 $cm^{-1}$ to 1386 $cm^{-1}$ and 1082 $cm^{-1}$ to 1074 $cm^{-1}$, respectively.

However, the most observable results of the frequency shift analysis, mainly due to the high absorbance and the clearly identified peak, was caused by the Amide II protein absorbance. The triple negative specimen red shifted in the spectra; our analysis showed the shift was 16 $cm^{-1}$, from 1546 $cm^{-1}$ to 1530 $cm^{-1}$. The Amide II

frequency shift was the most distinguishable; the absorbance peak also displayed a frequency shift in a visual observation of the spectra in Section 3.4.1.5.3. The Amide II frequency shift was the most identifiable peak shift throughout the frequency shift analysis with all our specimens.

## 3.5. Discussion / Conclusion

Mid-infrared spectroscopy has demonstrated an ability to provide biological information for breast tissues that is inaccessible or more difficult to acquire with other techniques. Daylight Solutions' Spero instrument provided vast improvements over the standard FTIR spectrometer in the speed with which we collected the data and in the quality of our data collection.

In this chapter, we collected spectral data on a number of breast tissue samples, both malignant and normal. We focused on spectral differences between ethnic groups, specifically white and black tissue samples, for both the normal and malignant specimens. We did not utilize other ethnic specimens due to small sample size. We also focused on the three main receptors typically utilized in breast cancer diagnosis and the location of the sample removed from the breast (lobular, ductal, etc.). Finally, during our normal and malignant breast comparisons, we also focused on the one triple negative specimen we obtained from CHTN.

We were able to focus and pinpoint our data collection techniques to exact locations on the breast samples by the following: (1) obtaining H & E stains of an

excised interior sample of each specimen and locating areas of interest, (2) taking an adjacent excised sample with which to collect spectral data on Spero, and (3) utilizing the mapping feature on Spero to match the location of interest previously defined on the H & E stain. The standard FTIR instrument would not have been able to collect such specific spectral data because of the inability to pinpoint data collection sites. The areas of fat cells on the breast samples do not have strong absorbance characteristics in the mid-infrared region, but the standard FTIR instrument has difficulty being able to differentiate between the fat cell and the other breast tissue regions on the samples to be able to collect valid absorbance spectral data.

We collected numerous images and spectra from each breast specimen; some of the examples of these data assemblages are shown in this chapter and in Appendix C and Appendix D. Because such a wide variety of histology diagnostic information was included with the tissue specimens, we focused our efforts on comparing ethnicities between normal and malignant samples, specimen extraction locations on malignant samples, and hormonal and protein receptors on malignant samples. For our preliminary analysis using Daylight Solutions' Spero microscope, we used the following three analysis techniques: intensity ratio parameter analysis, area-under-the-curve analysis, and frequency shift analysis.

Our preliminary results enabled us to differentiate between black and white normal breast tissue samples by analyzing the spectral data sets using intensity ratio parameter analysis. This initial analysis proved effective with our data at

wavenumbers 1030 $cm^{-1}$ and 962 $cm^{-1}$ with both Amide I and Amide II as the denominator parameter. However, because the six other ratios in the analysis set were not significant, calculating the area-under-the-curve for each spectrum provided further confirmation with our preliminary results in the normal white and black breast tissue differentiation. The mean spectra of the black normal breast specimens visibly had clear lower absorbance than the white specimens throughout the entire mid-infrared spectral range investigated in our preliminary research; thus, the AUC initial analysis results were expected and proved to be the most effective analysis technique in this case. Frequency shift analysis did not allow us to discern between white and black normal samples due to the low resolution of our data sampling and the close similarity of the mean spectra.

Preliminary results show we may also be able to identify the black and white malignant specimens via all three analysis techniques. Visible inspection of the spectra reveal significantly lower absorbance for the black malignant specimens throughout the entire spectrum evaluated, and the most effective analysis tool was the AUC because the delta between the two areas was so significant. In addition, all ratios we evaluated during our initial intensity ratio parameter analysis distinguished the black and white specimens with the exception of the Amide I/Amide II ratio and the Amide II/Amide I ratio. Finally, although less effective than the other two analysis techniques, frequency shift analysis also provided tentative results with two frequencies that allowed us to differentiate. Clear distinguishing

marks between white and black malignant samples were identified throughout all our preliminary analysis.

We were able to preliminarily identify the different types of malignant breast tissue (ductal, lobular, or phylloides) when we analyzed the spectral data results. We were successful classifying the three tissue types with our initial intensity ratio parameter analysis. In fact, with the Amide I absorbance peak as the denominator parameter, we identified one ratio per tissue type that specifically identified that tissue from the other two. In other words, the 1458/1654 ratio isolated phylloides tissue, the 1334/1654 ratio separated ductal samples, and 962/1654 identified lobular samples from the other two. Further information was available with this analysis that we identified in Section 3.4.1.2, but the most important results are summarized here.

Visual observation of the preliminary malignant breast spectra also determined the ductal tissue spectrum had lower absorbance that the lobular and phylloides spectra, and AUC analysis gave us potential delineation of the ductal tissues from both the lobular and phylloides specimens. The phylloides tissue samples' spectra had slightly higher absorbance throughout the mid-IR spectra range we investigated with the exception of the collagen and glycogen frequencies, where the absorbance was less than the other two tissue types. However, we were not able to distinguish the lobular and phylloides tissues from each other using AUC analysis. Likewise, frequency shift analysis allowed us to potentially identify phylloides tissue from lobular and ductal, but we were not able to isolate the ductal

and lobular specimens without implementing multiple wavenumbers. Thus, intensity ratio parameter analysis proved to be the best form of analysis which we utilized to preliminarily evaluate breast tissue type.

The intensity ratio parameter analysis proved to be a very effective tool for receptor identification. We evaluated the following three tissue receptor groups: ER+/PR+/HER2-, ER+/PR-/HER2-, and ER-/PR-/HER2-. Although the ratios with the denominator parameter utilizing either Amide I or Amide II had similar results, the delta between the ratios was much greater for Amide I ratios than Amide II ratios. When we utilized the weak absorbance peptide peak at 1334 $cm^{-1}$ in the numerator parameter, our preliminary results show we may be able to successfully distinguish all three receptor groups from each other. Several other ratios contained further receptor information, but this peptide ratio allowed all three receptor groups to be distinguished.

As we expected due to the similarity of the spectra, the AUC analysis did not prove to be an effective analysis tool to identify ER+/PR+/HER2-, ER+/PR-/HER2-, and ER-/PR-/HER2- specimens. The AUC for the three spectra did not have significant differences to allow us to make any distinction. Although frequency shift analysis did provide more useful analysis than AUC, it did not prove to be nearly as effective as intensity ratio parameter analysis. We were able to potentially classify both ER+/PR-/HER2- and ER-/PR-/HER2-, but there was no single frequency shift that allowed us to categorize ER+/PR+/HER2- specifically from the other two receptor specimens.

Interestingly, our initial results allowed us to use both the symmetrical and asymmetrical bending of the methyl protein and lipids frequencies (1398 $cm^{-1}$ and 1458 $cm^{-1}$, respectively) to isolate the ER-/PR-/HER2- specimen from the other two receptor groups. Although we only had one triple negative specimen, these results do provide preliminary results that indicate both these methyl protein and lipids frequencies may be influenced by the positive estrogen receptor. A larger sample size with a higher resolution sampling rate needs to be evaluated to confirm this analysis.

Finally, the preliminary results show we may be able to discern between normal and malignant samples. Specifically, we evaluated the following collection of specimens: normal white and malignant white; normal black and malignant black; aggregate normal and aggregate malignant specimens; and triple negative, aggregate normal, and malignant specimens excluding triple negative.

The intensity ratio parameter analysis was the most definitive tool we utilized in classifying the white and black samples as normal or malignant. There appears to be no significant difference between Amide I or Amide II as the denominator parameter. We identified six potential ratios that allow us to categorize the white specimens and seven potential ratios that allow us to categorize the black specimens as either normal or malignant. The only ratios that we were not able to utilize to preliminarily identify the normal or malignant white specimens were the Amide I/Amide II ratio and the symmetrical bending of methyl protein and lipid ratio; the only ratio we were not able to utilize to preliminarily

identify the normal or malignant black specimens was the nucleic acid and collagen band around 1080 $cm^{-1}$.

Because the spectrum for the normal white breast tissues in our initial results had slightly lower absorbance at the Amide III peak and the asymmetrical bending of methyl protein and lipid peak, we also expected the symmetrical bending of methyl protein and lipid to have a significant difference in the ratio between normal and malignant tissues. However, this was the one ratio besides the typical Amide I / Amide II ratio that may not satisfy the $p < .05$ significance factor. Likewise, because the collagen and glycogen bands tended to have useful ratio differences between our data sets, we expected the black specimens' collagen bands to display significant ratio differences between the malignant and normal samples. Additionally, the absorbance of the black malignant specimens was much lower throughout the spectra, including at Amide I and Amide II, except in the collagen and glycogen bands, so we expected the ratio to be useful in our normal and malignant categorization. However, the preliminary results did not match our analysis expectations.

Our initial AUC analysis may show significant differences between normal black and malignant black specimens; it also may allow us to identify normal white and malignant black specimens. However, we were unable to successfully separate normal white and malignant white samples or malignant white and normal black specimens. Similarly, our preliminary frequency shift analysis only allowed us to partially categorize the specimens. We were able to use frequency shifts at the

Amide II peak and both the symmetrical and asymmetrical bending of the methyl protein and lipids absorbance peaks to separate normal white and malignant white specimens. We were also able to utilize the methyl protein and the symmetrical bending of the methyl protein and lipid absorbance peaks to separate the normal and malignant black samples. No other frequencies exhibited a significant shift in our preliminary data, though, and we deemed this analysis form less effective than the intensity ratio parameter analysis.

When we evaluated all our normal and malignant breast specimens, we observed the normal spectrum in our preliminary data to have a higher absorbance throughout the spectra except in the collagen and glycogen frequencies. The Amide III and the asymmetrical bending of the methyl protein and lipid absorbance peaks specifically appear to be noticeably higher for the normal tissue. The greater absorbance shifted from the normal tissue to the malignant tissue at 1185 $cm^{-1}$, between the Amide III and collagen bands. The spectra also had a significantly visible frequency shift at the Amide II and the asymmetrical bending of the methyl protein and lipid absorbance peaks.

The most effective analysis tool we used with our preliminary data may be intensity ratio parameter analysis. When we utilized Amide I as the denominator parameter, we identified six ratios that had significant difference between the normal and malignant specimens and allowed us to categorize the samples; additionally, the differences were much greater than the analysis sets with Amide II in the denominator. Only the Amide II and the symmetrical bending of the methyl

protein and lipid absorbance peaks did not have significant differences between the normal and malignant ratios.

The area-under-the-curve technique did not seem to provide significant analysis with our preliminary data set; there was no significant difference between the area values that allowed us to categorize samples as normal or malignant. We also looked preliminarily at frequency shifts between the normal and malignant specimens; we identified potential frequency shifts in the Amide II peak and both the symmetrical and asymmetrical bending of the methyl protein and lipids absorbance peaks. Although the resolution of our data collection was at 4 $cm^{-1}$, the shifts specifically at Amide II and the asymmetrical bending have consistently been present in our preliminary data, and we anticipate this to be a viable analysis tool that could be consistently utilized in normal and malignant categorization of tissue samples in conjunction with intensity ratio parameter analysis.

After we evaluated all the normal and malignant specimens, we removed the one triple negative specimen from the malignant data and compared this specimen against the rest of the data. The difficulty in diagnosing triple negative tissue samples underscores the importance of finding alternative ways to diagnose these specimens as malignant. Our preliminary results indicate we may be able to distinguish the one triple negative specimen we had from the remaining malignant specimens and, more importantly, from the normal specimens by utilizing both intensity ratio parameter analysis and the preliminarily frequency shift analysis.

Area-under-the-curve analysis did not appear to be an effective technique with our preliminary data set to provide significant results.

Again, intensity peak ratio analysis with Amide I as the denominator parameter may be the most effective technique which we evaluated from our preliminary results. We identified five potential ratios we used to extricate the triple negative sample from the normal specimens. Similar to the other malignant specimens, the spectrum of the triple negative sample had a lower absorbance than the normal specimens throughout the spectra except in the collagen and glycogen regions, which had a higher absorbance. Thus, as expected from the spectra, the preliminary results show we may be able to discriminate using all three of the ratios evaluated in this region (1078 $cm^{-1}$, 1030 $cm^{-1}$, and 962 $cm^{-1}$) that showed possible significant differences between the normal and triple negative specimens. The other two ratios that may allow us to identify the triple negative specimen as malignant are the asymmetrical bending of the methyl protein and lipids at 1450 $cm^{-1}$ and the peptide group at 1334 $cm^{-1}$.

We identified three frequencies from our preliminary analysis results that may show a shift between the normal and triple negative specimens. Surprisingly, only two of the frequency shifts were the same as the shifts seen in the normal and malignant comparisons. The two same frequencies, Amide II and the symmetrical bending of the methyl protein and lipids, showed a greater shift between the normal and triple negative specimens than the shift between the normal and malignant

specimens, particularly at the Amide II absorbance peak, which had a red shift of 16 $cm^{-1}$.

Although AUC and frequency shift analysis provided useful results in our preliminary breast tissue analysis, intensity ratio parameter analysis may prove to be the most effective and consistent analysis technique. In unique situations, such as results dealing with malignant black tissue samples, AUC analysis may be more effective. Frequency shift analysis also shows promise as being a useful technique, but more experiments with higher resolution need to occur. Specifically, because one of the bands with a consistently significant frequency shift is the highly absorbent Amide II peak, this is a technique that could prove quite useful in separating the normal and malignant samples. Often, while researchers are using FTIR spectrometers, 4 $cm^{-1}$ seems to be a standard resolution during data collection. Any higher resolution would extend the already long data collection process. Utilizing Spero's tools could specify precise wavenumbers and/or bands while increasing the resolution of data points to 1 $cm^{-1}$ or less; the data collection time would be on the order of minutes instead of several hours.

More experiments with a larger sample size need to be conducted in order to substantiate these preliminary results, which should also include specific specimens with defined histology diagnostic characteristics. Multivariate analysis techniques could be considered in order to tease out more subtle information available in the spectra. Our work with the Spero quantum cascade laser microscope demonstrates great advantages over the standard FTIR spectrometer and is the first step in finally

making great strides in the mid-infrared cancer detection field that has been stagnant for so many years. With further exploitation of Spero's advantages, the speed of data collection and analysis can be improved ten-fold.

# Chapter 4

# Future Directions

Newer diagnostic capabilities for diseases include advances in technologies utilizing mid-infrared spectroscopy, specifically FTIR spectroscopy. However, inconsistencies in sample preparation and FTIR spectrometer processes, as well as the significant time required for spectral data acquisition, hinder the practical clinical applications of using mid-IR technologies. Tunable quantum cascade laser (QCL) microscopes possess potential solutions to overcome these inconsistencies and technical limitations, primarily due to high-resolution imaging over a wide field-of-view and rapid live-mode imaging. Taking advantage of tunable QCL microscope features to analyze and characterize diseased specimens and samples in the mid-IR spectrum have the potential to improve diagnostic accuracies, increase the speed of diagnostics, and decrease the medical time, effort, and costs of diagnostics.

We initially utilized standard FTIR spectroscopy techniques primarily to evaluate normal and malignant cell samples; we then evaluated QCL microscope instrumentation to investigate tissue specimens. We have developed preliminary analysis results that may characterize the various aspects of normal and diseased tissue specimens, including characterizing histological information included from the tissue bank. These initial results may indicate an ability to discriminate between the categories we evaluated, including the primary goal of distinguishing between the normal and diseased specimens. The ability to specify a precise target area on a high resolution image of the sample where we want to perform the spectral data collection is beneficial to enable the break-down and isolation of the specimens' unique differences to more completely characterize the histological information available.

In Chapter 2, we utilized Spero, a tunable QCL microscope, to investigate cervical tissue samples and develop preliminary results for the subgroups we created. We differentiated between normal and malignant samples; endocervix and ectocervix samples; epithelium, basal, and lamina propria layers; white and black specimens; and malignant squamous cell carcinoma and glassy cell carcinoma. To evaluate these subgroups, we used the following three analysis techniques, all of which may prove to be successful techniques for future work: intensity ratio parameter analysis, area-under-the-curve analysis, and frequency shift analysis. Our initial data collection revealed the importance of being able to acquire the high resolution images real time and expedite the spectra data collection. By being able

to pinpoint on the sample specifically what exact area of the tissue we wanted to analyze, our preliminary results indicate we may be able to then discriminate between the subgroups we were evaluating. FTIR spectrometers would not have allowed us to identify the samples to this specificity.

In Chapter 3, we expounded on our cervical work developed in Chapter 2 and utilized similar techniques for our breast specimens. The preliminary results we developed may show the ability to discriminate normal and malignant specimens; white and black specimens; ductal, lobular, and phylloides malignant samples; and hormonal and protein receptor variances, including distinguishing triple negative samples from normal samples. We again used the same three analysis techniques; although all three techniques provided useful conclusions, the intensity ratio parameter analysis was the most succinct and beneficial in our preliminary determinations. The accuracy and precision benefits of a tunable QCL microscope were essential in our initial breast tissue analysis due to the low mid-IR absorbance and reflectance characteristics of the high fat cell content of the specimens. Not having the capability to specify the tissue portion of the samples would have made the spectra acquisitions extremely difficult; FTIR spectrometers would not have allowed us to identify the samples to this specificity.

Future work should expound on the preliminary analysis we developed by using larger sample sizes in the experiments. We were limited by the number of applicable tissue specimens we had for our research, and having larger sample sizes is the next step to develop, verify, and validate new algorithms for use with this

instrumentation and these techniques. In addition, obtaining more histological information with each specimen would expand the mid-IR data characterization sets, leading to further environmental and biographical factors to explore. Further work could be performed to identify spectral characteristics of the individual breast receptors, which would further benefit the triple negative diagnostics. More complete histological information with each specimen could also allow spectra comparison between percentages of tumor on similar tissue samples.

Other methods of the same analysis techniques can also be explored, such as using reflectance peaks in place of absorbance peaks in intensity ratio parameter analysis and frequency shift analysis. Knowing specific wavenumbers of interest in the intensity ratio parameter analysis could greatly increase the speed of the data collection by identifying specific wavenumbers at which to collect the data instead of acquiring the full spectrum of absorbance data; one problem that would have to be overcome, however, would be taking into account the red or blue frequency shifts between sample subsets. Furthermore, a higher resolution of the data collection (1 $cm^{-1}$ as opposed to 4 $cm^{-1}$) would greatly increase the effectiveness of frequency shift analysis. Because the resolution of the QCL microscope is high enough, future work could also investigate the spectral characteristics and biomarkers of individual cells either on the tissue samples themselves or by using incubated cell samples. Moreover, area-under-the-curve analysis could be utilized in smaller wavenumber ranges, such as under individual absorbance or reflectance peaks.

In addition to increased sample size, one of the most definitive areas for future work is exploring more powerful analysis techniques that may allow greater concrete results. Utilizing a combination of partial least squares regression analysis and linear discriminate analysis techniques may be able to tease out additional information that is hidden in the spectra and not readily accessible through the three techniques we explored. Principal component analysis could also be evaluated, although partial least squares analysis tends to be more powerful, especially when used in conjunction with linear discriminate analysis.

In this thesis, we established effective techniques of differentiating normal and malignant tissue specimens. We demonstrated we could achieve valid results in a more consistent manner than previous mid-IR techniques, and that the data collection methods via QCL microscopy could be completed orders of magnitude faster than prior methods. We have also demonstrated that due to the high resolution and real-time imaging capabilities of a QCL microscope, we can vastly expand the spectral characterizations of biological and environmental characteristics of the specimens being evaluated. This thesis highlights the benefits of moving on from FTIR spectroscopy to QCL microscopy in mid-IR imaging and data analysis techniques.

Ultimately, we want cheaper/better/faster results that have clinical significance, but more work is needed to investigate if that technology is achievable. FTIR spectroscopy has been a leading technology used in mid-IR imaging and analysis, but utilizing current FTIR spectrometers has not made great strides in the

last decade and appears to have reached close to the maximum effectiveness this technology can achieve without improvements in the instrumentation. This dissertation has demonstrated that QCL microscopy technology may be a big stride in the direction of dramatic improvements to mid-IR imaging and analysis capabilities.

# References


1. Hoyert, Donna L., et al. "Deaths: Preliminary Data for 2011," *National Vital Statistics Reports:* vol 61, no 6, 2012.

2. Heron, Melonie. "Deaths: Leading Causes for 2009," *National Vital Statistics Reports:* vol 61, no 7, 2012.

3. Movasaghi, Zanyar, Shazza Rehman, and Dr Ihtesham ur Rehman. "Fourier transform infrared (FTIR) spectroscopy of biological tissues." *Applied Spectroscopy Reviews* 43.2 (2008): 134-179.

4. Kelly, Jemma G., et al. "A spectral phenotype of oncogenic human papillomavirus-infected exfoliative cervical cytology distinguishes women based on age." *Clinica chimica acta* 411.15 (2010): 1027-1033.

5. Machado, M., J. Hadgraft, and M. E. Lane. "Assessment of the variation of skin barrier function with anatomic site, age, gender and ethnicity." *International journal of cosmetic science* 32.6 (2010): 397-409.

6. Shai, Yechiel. "ATR-FTIR Studies in Pore Forming and Membrane Induced Fusion Peptides." *Biochimica et Biophysica Acta (BBA)-Biomembranes* (2012).

7. Walsh, Michael J., et al. "ATR microspectroscopy with multivariate analysis segregates grades of exfoliative cervical cytology." *Biochemical and biophysical research communications* 352.1 (2007): 213-219.

8. Walsh, Michael J., et al. "Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology." *Vibrational Spectroscopy* 60 (2012): 23-28.

9. Lane, Randy, and Seong S. Seo. "Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy Method to Differentiate Between Normal and Cancerous Breast Cells." *Journal of Nanoscience and Nanotechnology* 12.9 (2012): 7395-7400.

10. Sahu, R. K., et al. "Can Fourier transform infrared spectroscopy at higher wavenumbers (mid IR) shed light on biomarkers for carcinogenesis in tissues?." *Journal of biomedical optics* 10.5 (2005): 054017-054017.

11. Sade, Asli, et al. "Celecoxib reduces fluidity and decreases metastatic potential of colon cancer cell lines irrespective of COX-2 expression." *Bioscience reports* 32.1 (2012): 35-44.

12. Wong, Patrick TT, et al. "Characterization of exfoliated cells and tissues from human endocervix and ectocervix by FTIR and ATR/FTIR spectroscopy." *Biospectroscopy* 1.5 (1995): 357-364.

13. Ozek, Nihal Simsek, et al. "Characterization of microRNA-125b expression in MCF7 breast cancer cells by ATR-FTIR spectroscopy." *Analyst* 135.12 (2010): 3094-3102.

14. Sun, Xiaoliang, et al. "Detection of lung cancer tissue by attenuated total reflection–Fourier transform infrared spectroscopy—a pilot study of 60 samples." *Journal of Surgical Research* (2012)

15. Fujioka, Naoko, et al. "Differences between infrared spectra of normal and neoplastic human gastric cells." *Spectroscopy: An International Journal* 18.1 (2004): 59-66.

16. Yano, Kazuyuki, et al. "Direct measurement of human lung cancerous and noncancerous tissues by Fourier transform infrared microscopy: can an infrared microscope be used as a clinical tool?." *Analytical Biochemistry* 287.2 (2000): 218-225.

17. Lewis, Paul D., et al. "Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum." *BMC cancer* 10.1 (2010): 640.

18. Mariangela, Cestelli-Guidi, et al. "Experimental ATR device for real-time FTIR imaging of living cells using brilliant synchrotron radiation sources." *Biotechnology advances* (2011).

19. Mostaço-Guidolin, Leila Büttner, et al. "Fourier transform infrared spectroscopy of skin cancer cells and tissues." *Applied Spectroscopy Reviews* 44.5 (2009): 438-455.

20. Maziak, Donna E., et al. "Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: an exploratory study." *Cancer detection and prevention* 31.3 (2007): 244-253.

21. Sulé-Suso, J., et al. "FT-IR microspectroscopy as a tool to assess lung cancer cells response to chemotherapy." *Vibrational spectroscopy* 38.1 (2005): 179-184.

22. Noreen, Razia, et al. "FTIR spectro-imaging of collagens for characterization and grading of gliomas." *Biotechnology advances* (2012).

23. Zwielly, A., et al. "Grading of intrinsic and acquired cisplatin-resistant human melanoma cell lines: an infrared ATR study." *European Biophysics Journal* 40.6 (2011): 795-804.

24. Nasse, Michael J., et al. "High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams." *Nature Methods* 8.5 (2011): 413-416.

25. Rigas, Basil, et al. "Human colorectal cancers display abnormal Fourier-transform infrared spectra." *Proceedings of the National Academy of Sciences* 87.20 (1990): 8140-8144.

26. Gazi, Ehsan, et al. "Imaging ToF-SIMS and synchrotron-based FT-IR microspectroscopic studies of prostate cancer cell lines." *Applied surface science* 231 (2004): 452-456.

27. German, Matthew J., et al. "Infrared spectroscopy with multivariate analysis potentially facilitates the segregation of different types of prostate cell." *Biophysical journal* 90.10 (2006): 3783-3795.

28. Taylor, S. E., et al. "Infrared spectroscopy with multivariate analysis to interrogate endometrial tissue: a novel and objective diagnostic approach." *British journal of cancer* 104.5 (2011): 790-797.

29. Yang, Wenyue, et al. "In situ evaluation of breast cancer cell growth with 3D ATR-FTIR spectroscopy." *Vibrational Spectroscopy* 49.1 (2009): 64-67.

30. Li, Qing-Bo, et al. "In vivo and in situ detection of colorectal cancer using Fourier transform infrared spectroscopy." *World journal of gastroenterology: WJG* 11.3 (2005): 327.

31. Mostaço-Guidolin, Leila Büttner, et al. "Molecular and chemical characterization by Fourier transform infrared spectroscopy of human breast cancer cells with estrogen receptor expressed and not expressed." *Journal of Spectroscopy* 24.5 (2010): 501-510.

32. Bogomolny, Evgeny, et al. "Monitoring of viral cancer progression using FTIR microscopy: a comparative study of intact cells and tissues." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1780.9 (2008): 1038-1046.

33. Di Giambattista, Lucia, et al. "New marker of tumor cell death revealed by ATR-FTIR spectroscopy." *Analytical and bioanalytical chemistry* 399.8 (2011): 2771-2778.

34. Lv, Jungang, et al. "Optical Observation of Lung Cancer With Attenuated Total Reflectance–Fourier Transform Infrared Microscope (ATR-FTIR) and Confocal Raman Microscope." *Spectroscopy Letters* 44.5 (2011): 312-317.

35. Llabjani, Valon, et al. "Polybrominated diphenyl ether-associated alterations in cell biochemistry as determined by attenuated total reflection Fourier-transform infrared spectroscopy: a comparison with DNA-reactive and/or endocrine-disrupting agents." *Environmental science & technology* 43.9 (2009): 3356-3364.

36. Mordechai, S., et al. "Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma." *Journal of microscopy* 215.1 (2004): 86-91.

37. Keith, Frances N., Rohith K. Reddy, and Rohit Bhargava. "Practical protocols for fast histopathology by Fourier transform infrared spectroscopic imaging." *Biomedical Optics (BiOS) 2008*. International Society for Optics and Photonics, 2008.

38. Bakry, R., et al. "Protein profiling for cancer biomarker discovery using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and infrared imaging: a review." *Analytica chimica acta* 690.1 (2011): 26-34.

39. Kabachkov, E., et al. "Research of photocatalytic degradation of HeLa cells at the TiO< sub> 2</sub> interface by ATR-FTIR and fluorescence microscopy." *Journal of Photochemistry and Photobiology A: Chemistry* 217.2 (2011): 425-429

40. Kelly, Jemma G., et al. "Robust classification of low-grade cervical cytology following analysis with ATR-FTIR spectroscopy and subsequent application of self-learning classifier eClass." *Analytical and bioanalytical chemistry* 398.5 (2010): 2191-2201.

41. Pounder, F. Nell, and Rohit Bhargava. "Spectral histology of breast tissue using mid-infrared spectroscopic imaging." *SPIE BiOS: Biomedical Optics*. International Society for Optics and Photonics, 2009.

42. Pijanka, Jacek Klaudiusz, et al. "Spectroscopic signatures of single, isolated cancer cell nuclei using synchrotron infrared microscopy." *Analyst* 134.6 (2009): 1176-1181.

43. Caine, Sally, et al. "The application of Fourier transform infrared microspectroscopy for the study of diseased central nervous system tissue." *Neuroimage* 59.4 (2012): 3624-3640.

44. American Cancer Society. "Cancer Facts and Figures." *http://www.cancer.org/research/cancerfactsstatistics.index*. American Cancer Society, Inc. (2015).

45. Centers for Disease Control and Prevention. "Deaths and Mortality." *http://www.cdc.gov/nchs/fastats/deaths.htm.* U.S. Department of Health and Human Services (2015).

46. Centers for Disease Control and Prevention. "Cervical Cancer Statistics." *http://www.cdc.gov/ cancer/cervical/statistics/.* U.S. Department of Health and Human Services (2014).

47. World Cancer Report 2014. "Media Centre Cancer Fact Sheet." *http://www.who.int/mediacentre/factsheets/fs297/en/.* World Health Organization (2015).

48. de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. The Lancet Oncology 2012;13: 607-615.

49. Daylight Solutions. "SPERO: Introducing the Industry's First Laser-based Infrared Microscopy Platform." *http://www.daylightsolutions.com/life-sciences/spero.htm.* (2015).

50. Rowlette, J. *et al.* High-confidence, high-throughput screening with high-def IR microspectroscopy. *BioOptics World* (March/April 2014).

51. Daylight Solutions. "SPERO: Product Brocure." *http://www.daylightsolutions.com/assets/005/5511.pdf.* (2015).

52. Cooperative Human Tissue Network Southern Division. Tissue Collection and Banking Facility. *http://www.uab.edu/medicine/tcbf/chtn-sd-sidebar.* University of Alabama Birmingham (n.d.).

53. International Crystal Laboratories. "Calcium Fluoride ($CaF_2$) Optical Crystals." *http://www.internationalcrystal.net/optics_06.htm* (2014).

54. American Cancer Society. "What is cancer of the cervix?" *http://www.cancer.org/cancer/cervicalcancer/overviewguide/cervical-cancer-overview-what-is-cervical-cancer*. American Cancer Society, Inc. (2014).

55. Tsuboi, M., Suzuki, M., Overman, S., & Thomas, G. "Intensity of the Polarized Raman Band at 1340-1345 $cm^{-1}$ as an Indicator of Protein a-Helix Orientation:

Application to Pf1 Filamentous Virous." *Biochemistry* 39(10), 2677–2684 (2000).

56. Gray, H. J., Garcia, R., Tamimi, H. K., Koh, W. J., Goff, B. A., Greer, B. E., & Paley, P. J. "Glassy cell carcinoma of the cervix revisited." *Gynecologic Oncology* 85(2), 274-277 (May 2009).

57. Deshpande, A. H., Kotwal, M. N., & Bobhate, S. K. "Glassy cell carcinoma of the uterine cervix a rare histology. Report of three cases with a review of the literature." *Indian Journal of Cancer 41(2) (2004).*

58. Jones, Oliver. "The Vagina: Histological Structure." *http://teachmeanatomy.info/pelvis/female-reproductive-tract/vagina/*. TeachMeAnatomy (2015).

59. The Human Protein Atlas. "Uterine, Cervix." *http://www.proteinatlas.org/learn/dictionary/normal/cervix,+uterine/detail+2*. SciLifeLab (n. d.).

60. Esque, Della. "Cervix." *HistoQuarterly* (Jan. 2013).

61. Breast Cancer Research Foundation*.* "Breast Cancer Statistics & Resources." *http://www.bcrfcure.org/breast-cancer-statistics-resources?gclid=CPPdt5Pwz8sCFQiqaQod8eUGXg*. (n. d.).

62. Ferlay, J*. et al.* "Breast Cancer Statistics." *http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics.* World Cancer Research Fund International (2014).

63. American Cancer Society. "Cancer Facts and Figures 2015." *http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2015/index*. American Cancer Society, Inc. (2015).

64. New Health Guide. "Breast Cancer Pictures." *http://www.newhealthguide.org/Breast-Cancer-Pictures.html*. (2014).

65. Tzou, K. S. "An Introduction to Breast Cancer: Biology, Pathology, and the Latest in Screening and Diagnostic Tools." *http://docplayer.net/3567681-An-introduction-to-breast-cancer-biology-pathology-and-the-latest-in-screening-and-diagnostic-tools.html.* Mayo Clinic Jacksonville, Department of Radiation Oncology (n. d.).

66. BreastCancer.Org. "Types of Breast Cancer." *http://www.breastcancer.org/symptoms/types* (2015).

67. BreastCancer.Org. "Phyllodes Tumors of the Breast." *http://www.breastcancer.org/symptoms/types/phyllodes* (2015).

68. Lannin, D. R., Konstantakos, A. K., Raaf, J. H., & Geibel, J. "Cystosarcoma Phyllodes." *Medscape* (Sept. 2014).

69. Grau, A. M., Chakravarthy, A. B., & Chugh, R. "Phyllodes tumors of the breast." *http://www.uptodate.com/contents/phyllodes-tumors-of-the-breast*. UpToDate (Feb. 2014).

70. Triple Negative Breast Cancer Foundation. "Understanding Triple Negative Breast Cancer." *http://www.tnbcfoundation.org/understanding-triple-negative-breast-cancer/* (2015).

71. BreastCancer.Org. "Triple Negative Breast Cancer." *http://www.breastcancer.org/symptoms/diagnosis/trip_neg* (Oct. 2015).

72. Susan G. Komen Foundation. "Triple Negative Breast Cancer." *https://ww5.komen.org/uploadedFiles/_Komen/Content/About_Breast_Cancer/Tools_and_Resources/Fact_Sheets_and_Breast_Self_Awareness_Cards/Triple%20Negative%20Breast%20Cancer.pdf* (2014).

73. National Breast Cancer Foundation. "Triple Negative Breast Cancer." *http://www.nationalbreastcancer.org/triple-negative-breast-cancer.* National Breast Cancer Foundation, Inc. (2014).

74. Johns Hopkins Medicine. "Triple Negative Breast Cancer." *http://www.hopkinsmedicine.org/breast_center/breast_cancers_other_conditions/triple_negative_breast_cancer.html.* The Johns Hopkins University (2015).

75. Sharma, P. *et al.* "Germline BRCA mutation evaluation in a prospective triple-negative breast cancer registry: implications for hereditary breast and/or ovarian cancer syndrome testing." *Breast Cancer Research and Treatment* 145(3), 707-714 (2014).

76. Stecklein, S. R. & Sharma, P. "Tumor homologous recombination deficiency assays: another step closer to clinical application?" *Breast Cancer Research and Treatment* 16(4), 409 (2014).

77. Bhalla, K. N., Sharma, P., Das Gupta, S., Chauhan, L, & Stecklein, S. "Treatment with histone deacetylase inhibitors creates 'BRCAness' and sensitizes human triple negative breast cancer cells to PARP inhibitors and cisplatin." *Breast Cancer Research and Treatment,* 72(24), S3-7 (2012).

78. Hudis, C. A. & Gianni, L. "Triple-Negative Breast Cancer: An Unmet Medical Need." *The Oncologist,* 16(1), 1-11 (2011).

79. Mello, M. L. S. & Vidal, B. C. "Changes in the Infrared Microspectroscopic Characteristics of DNA Caused by Cationic Elements, Different Base Richness and Single-Stranded Form." *PLoS One*, Public Library of Science. 7(8), (Aug. 2012).

80. Ratner, B. D., Hoffman, A. S., Schoen, F. J., & Lemons, J. E. *Biomaterials Science: An Introduction to Materials in Medicine.* Waltham: Academic Press, 3rd ed., 2013. Print.

# Appendix A

# Cervical Normal

### A1.04-10-A245a

| | |
|---|---|
| Unique ID | 04-10-A245a |
| Age | 32 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | none |

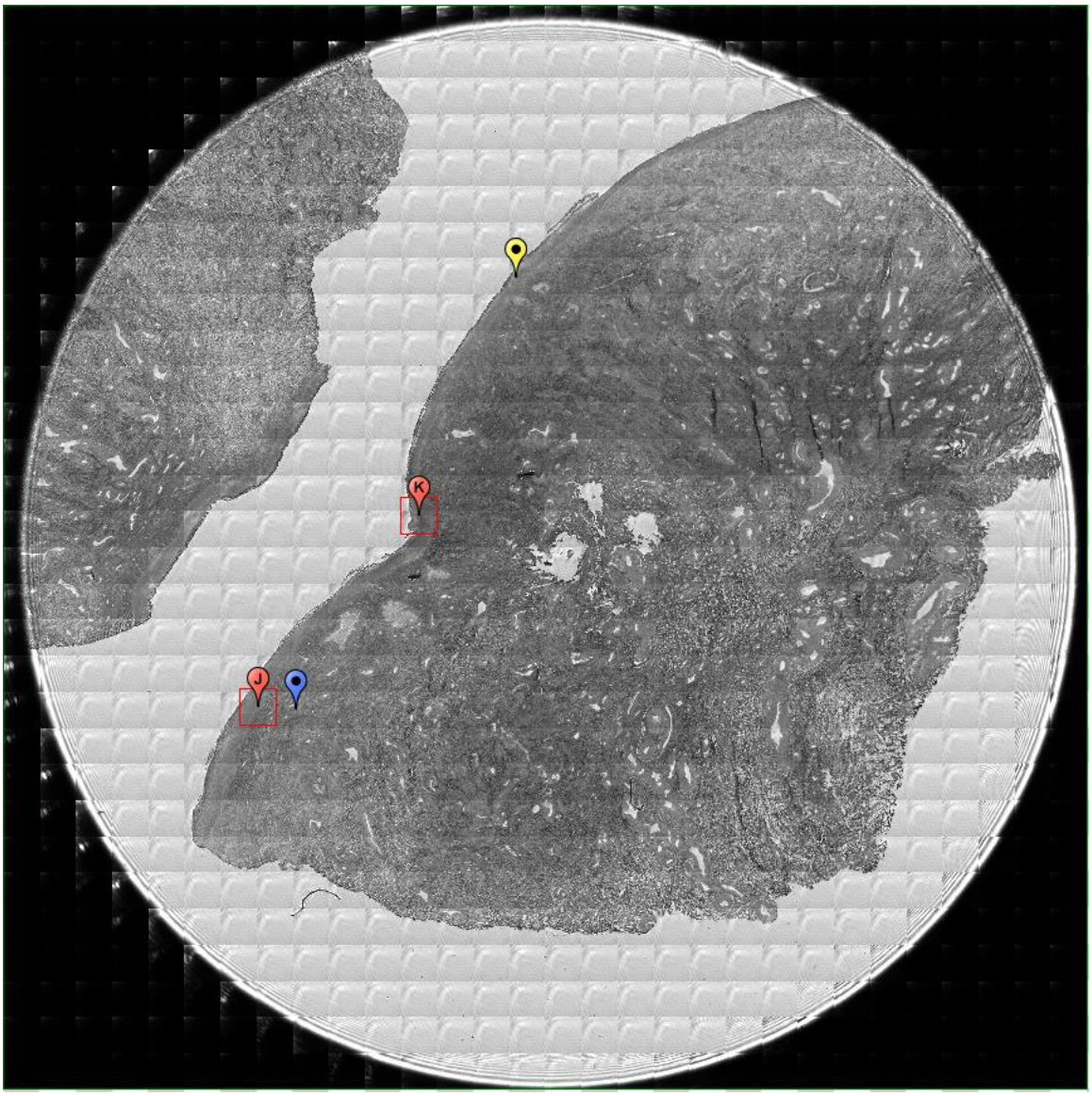


**Figure 0.1. Specimen 04-10-A245a, normal cervical tissue.**

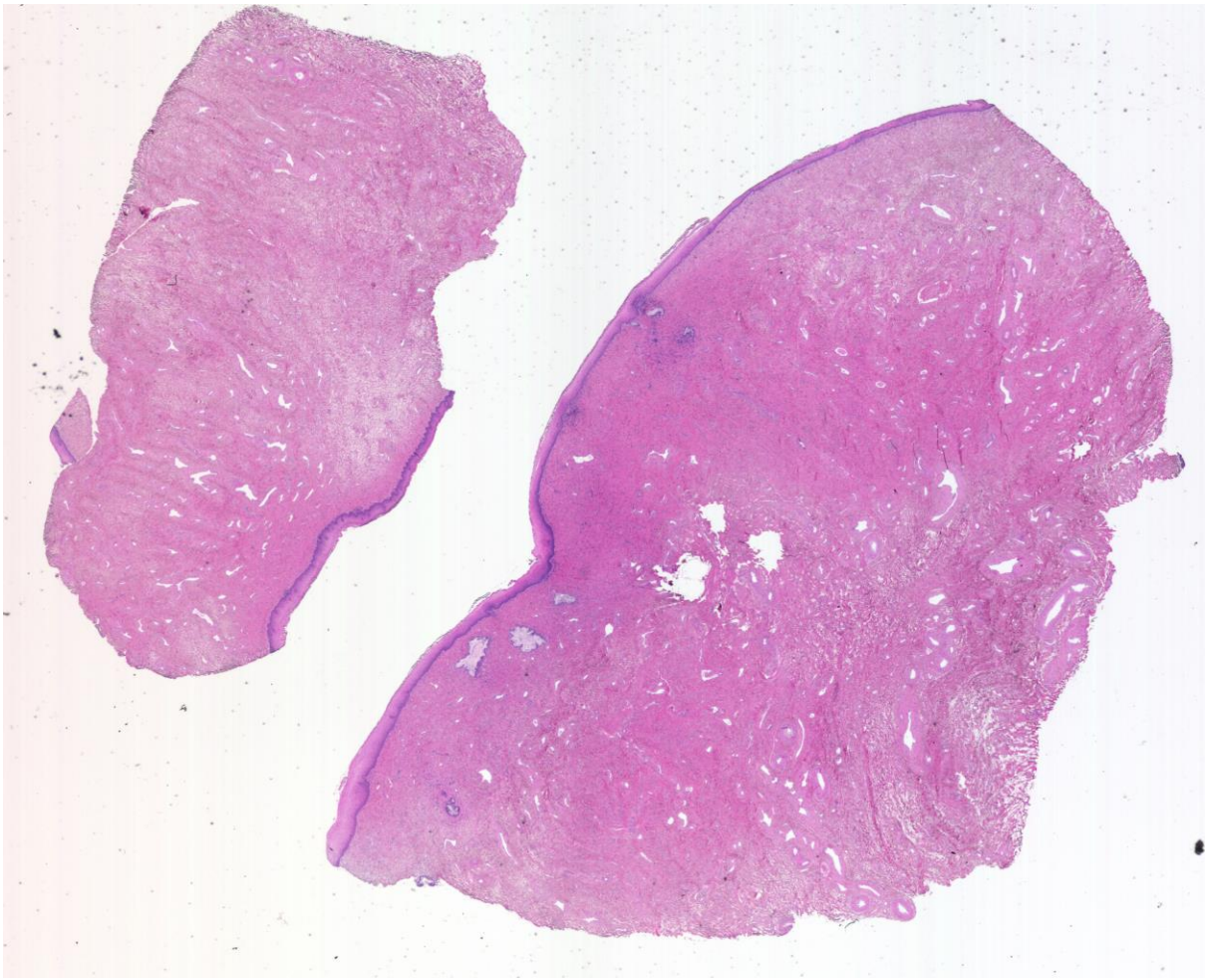


**Figure 0.2. Specimen 04-10-A245a, H&E stain, normal cervical tissue.**

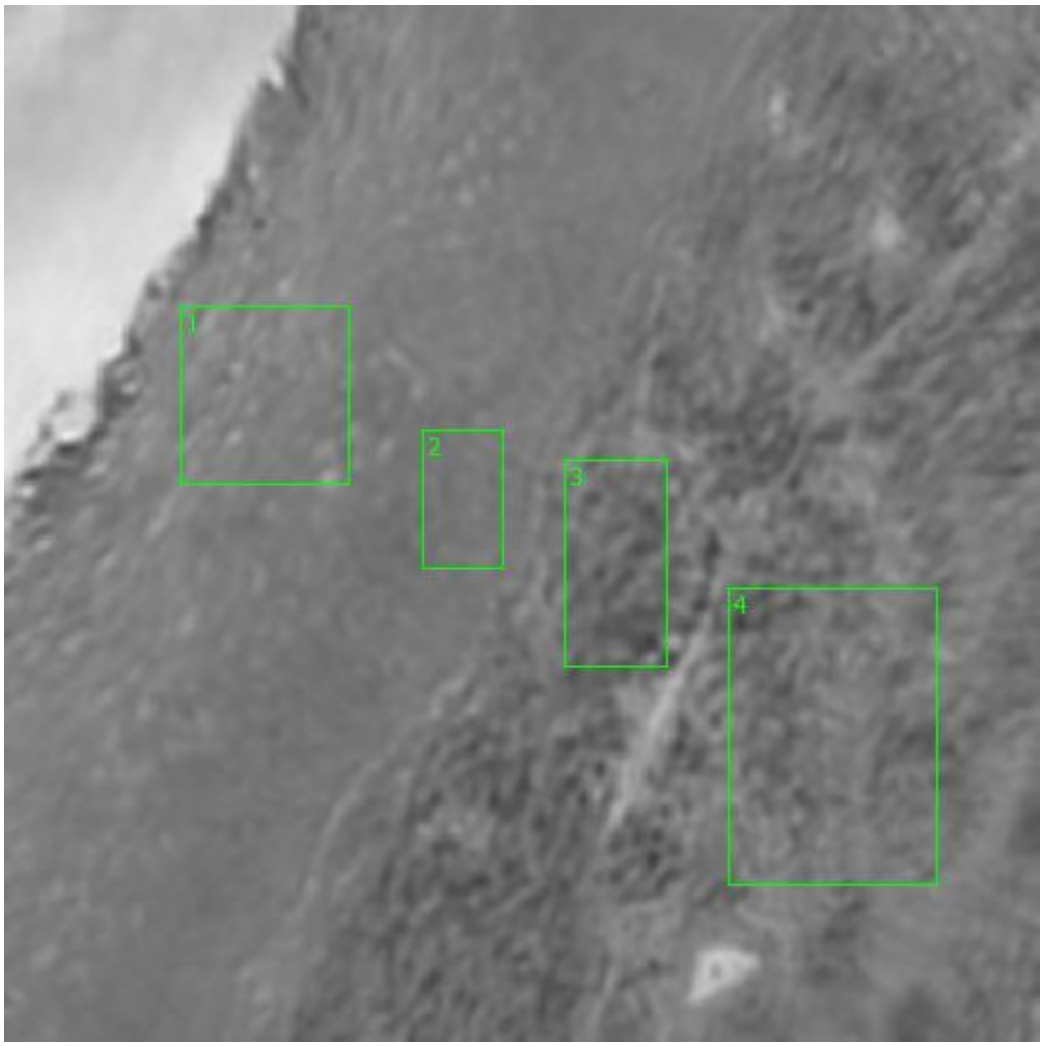



**Figure 0.3. Point J Regions of Interest**
**Region 1: stratified squamous epithelium**
**Region 2: basal layer**
**Regions 3, 4: basement membrane, lamina propia**

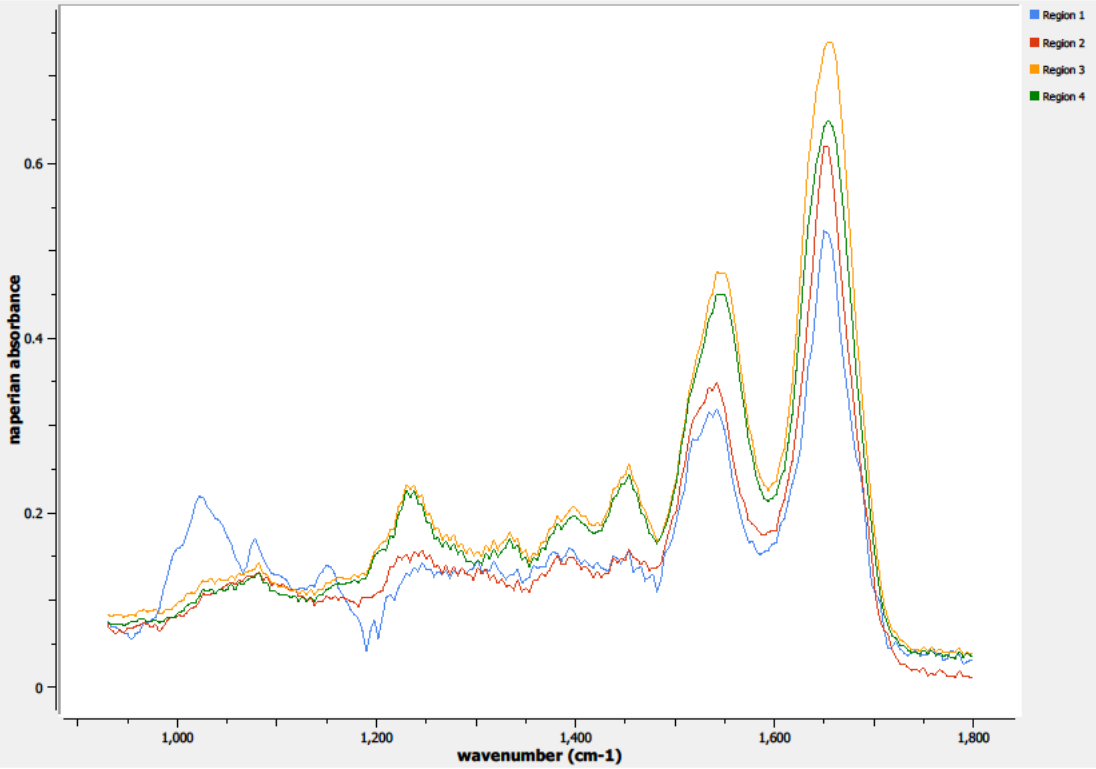



**Figure 0.4. Point J Spectra**

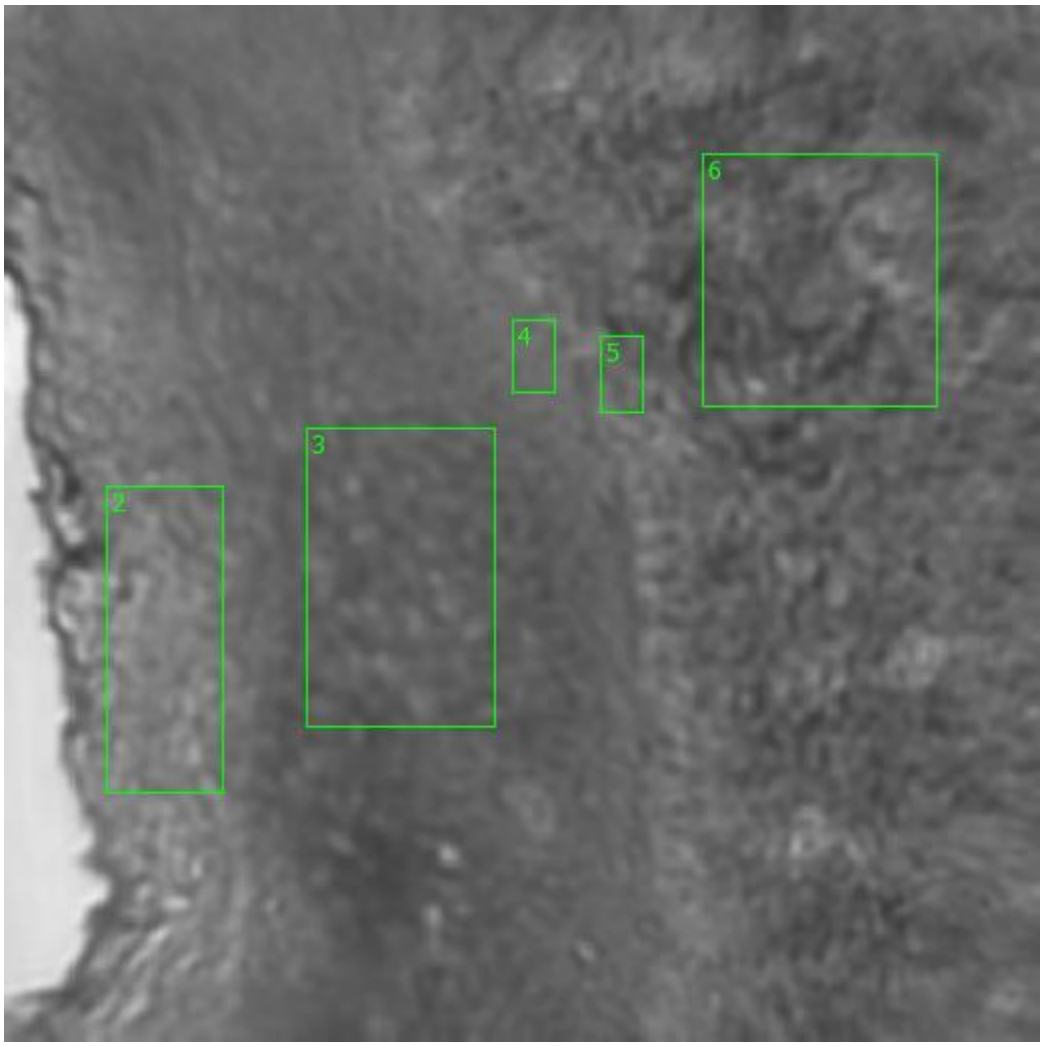



**Figure 0.5. Point K Regions of Interest**
**Region 2: stratified squamous epithelium**
**Region 3: columnar epithelium**
**Region 4: basal layer**
**Regions 5, 6: basement membrane, lamina propia**

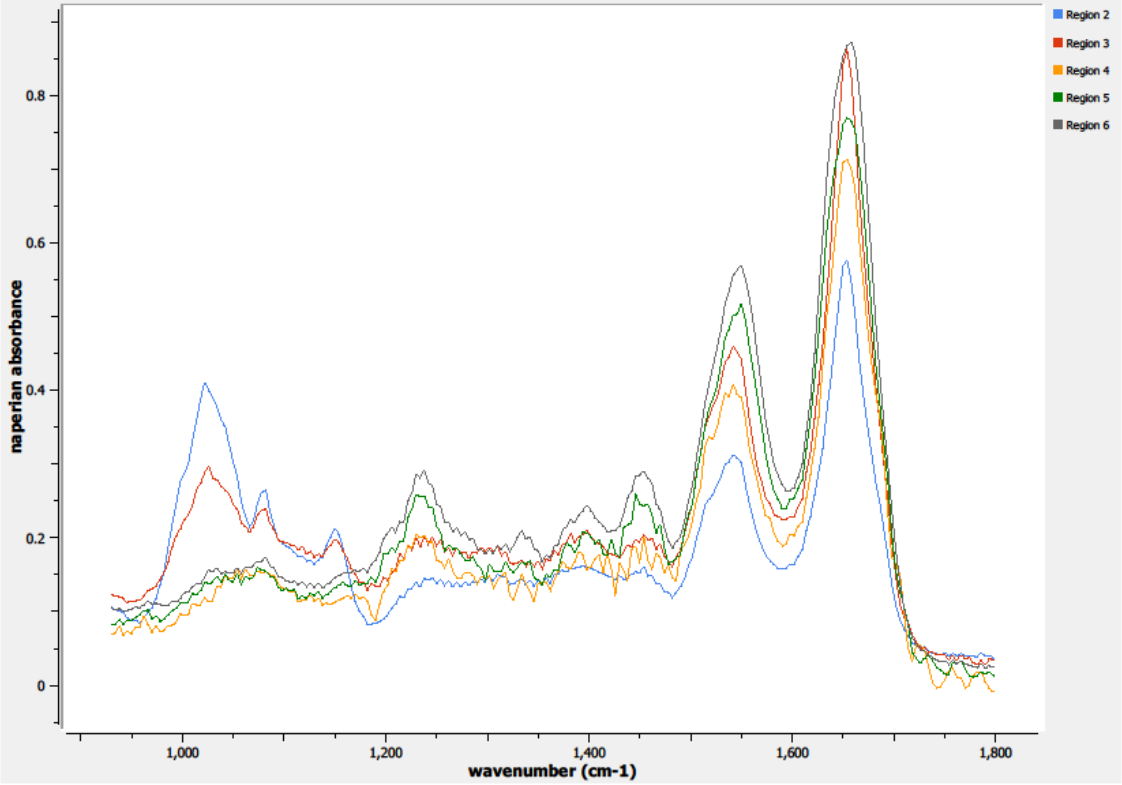



**Figure 0.6. Point K Spectra**

## A2.06-04-A113a

| | |
|---|---|
| Unique ID | 06-04-A113a |
| Age | 56 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

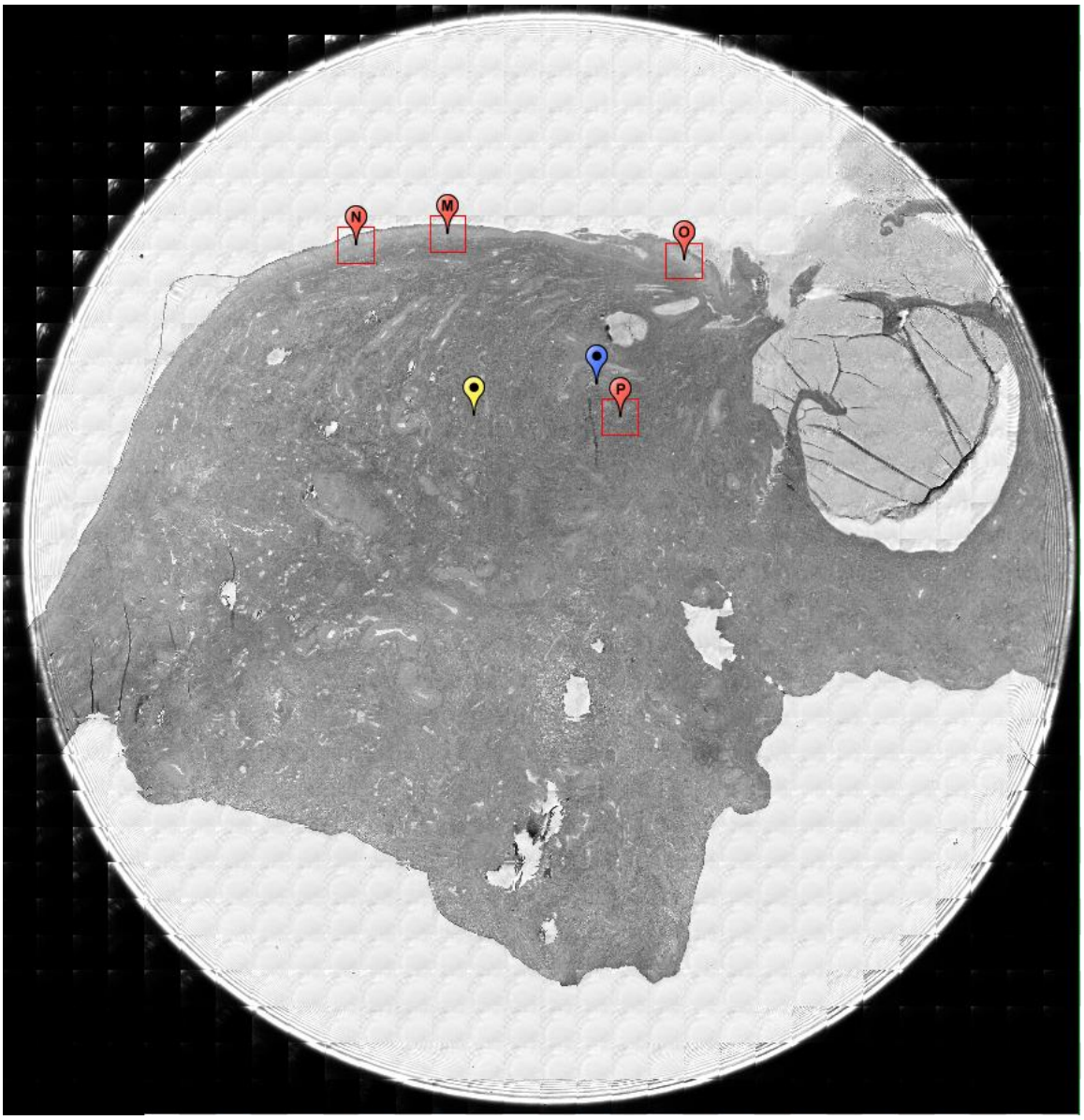



**Figure 0.7. Specimen 06-04-A113a, normal cervical tissue.**

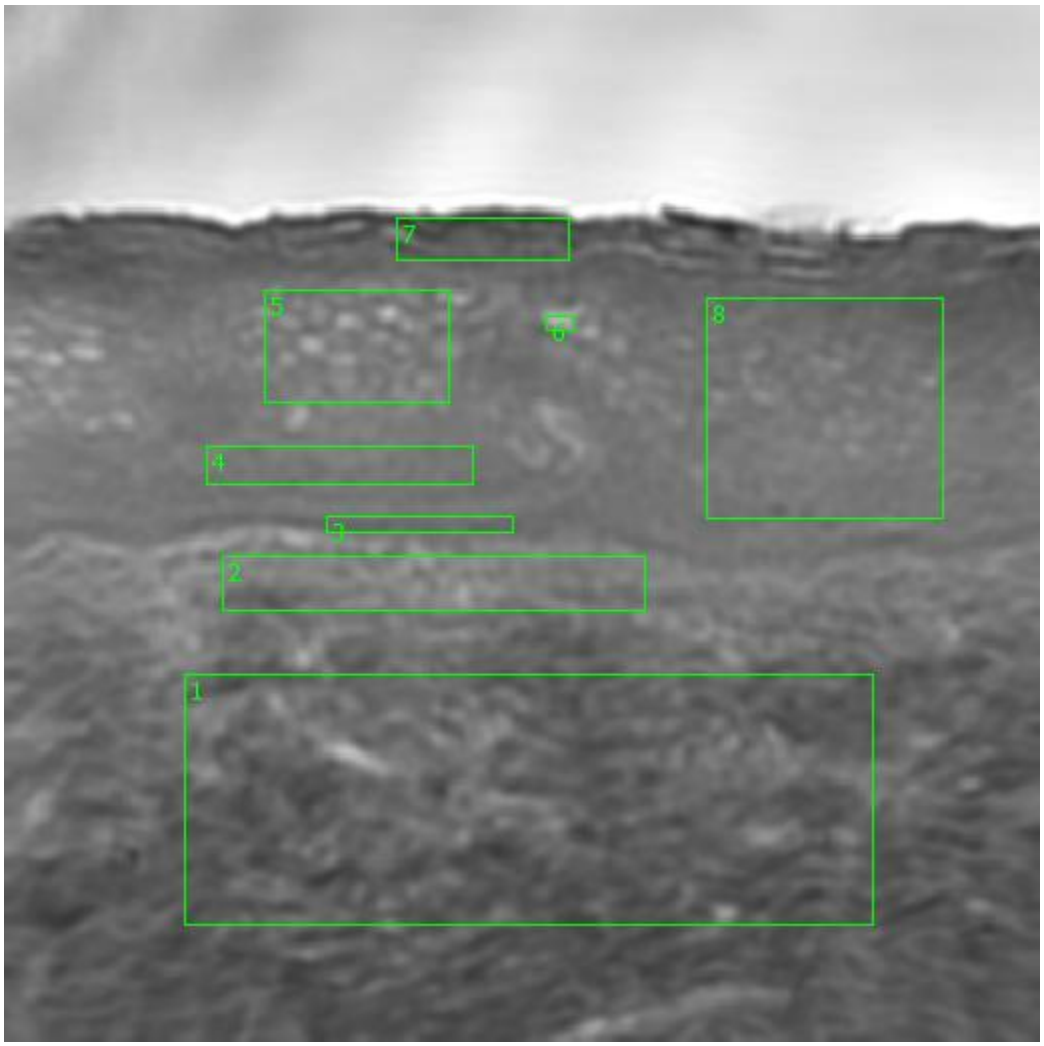



**Figure 0.8. Point M Regions of Interest**
**Regions 1, 2: basement membrane, lamina propia**
**Region 3: basal layer**
**Region 4: parabasal layer**
**Regions 5, 8: columnar epithelium**
**Region 6: individual columnar epithelial cell**
**Region 7: stratified squamous epithelium, epithelial surface**

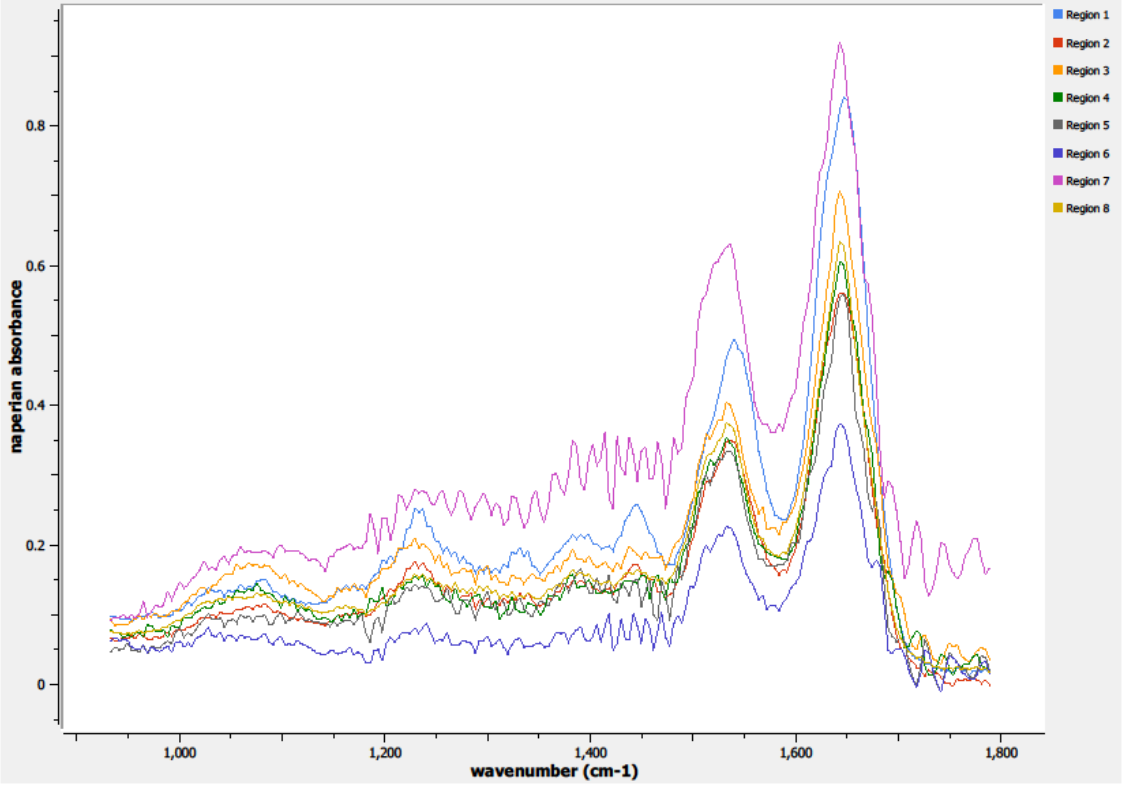



**Figure 0.9. Point M Spectra**

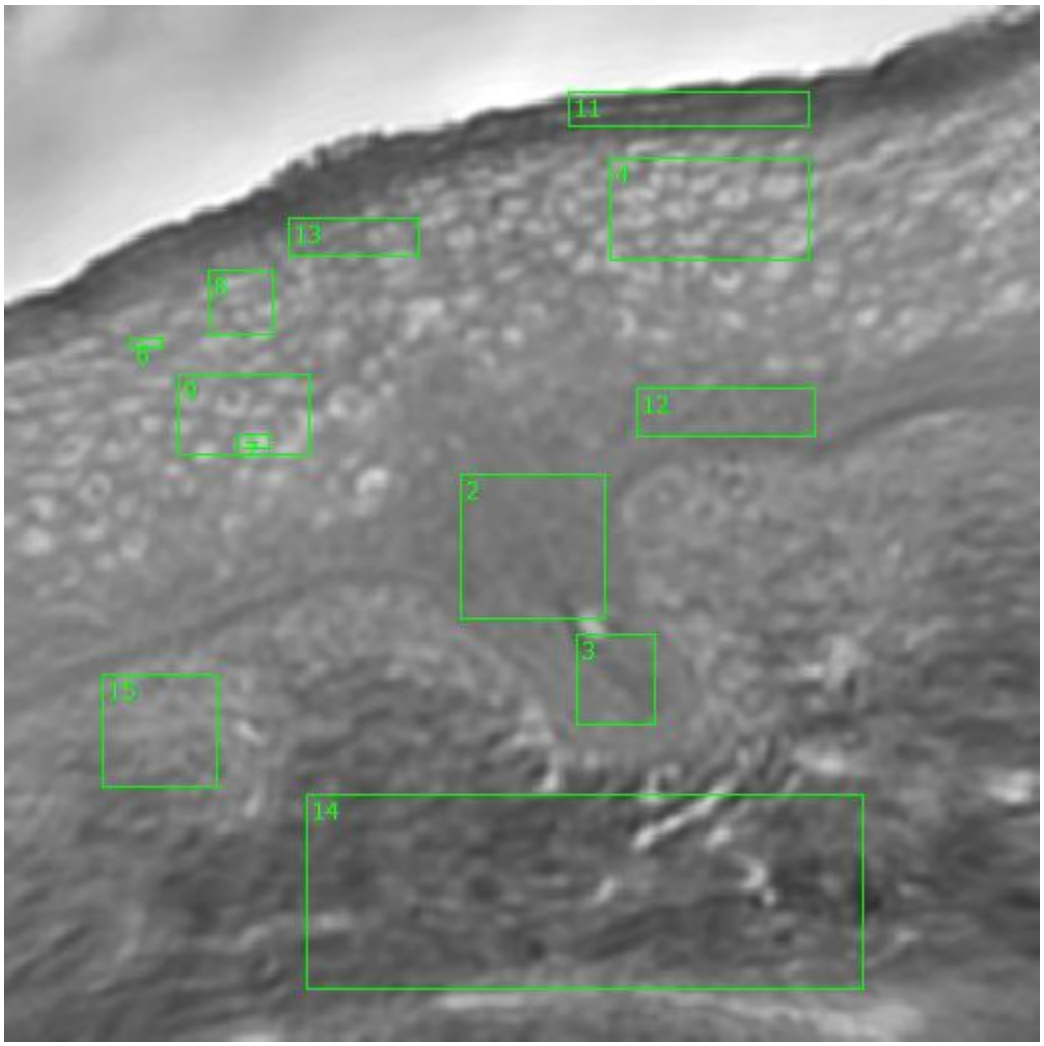



**Figure 0.10. Point N Regions of Interest**
**Regions 2, 3, 12: basal layer**
**Regions 4, 9: columnar epithelium**
**Regions 8, 13: stratified squamous epithelium**
**Region 6: individual squamous epithelial cell**
**Region 7: individual columnar epithelial cell**
**Region 11: stratified squamous epithelium, epithelial surface**
**Regions 14, 15: basement membrane**

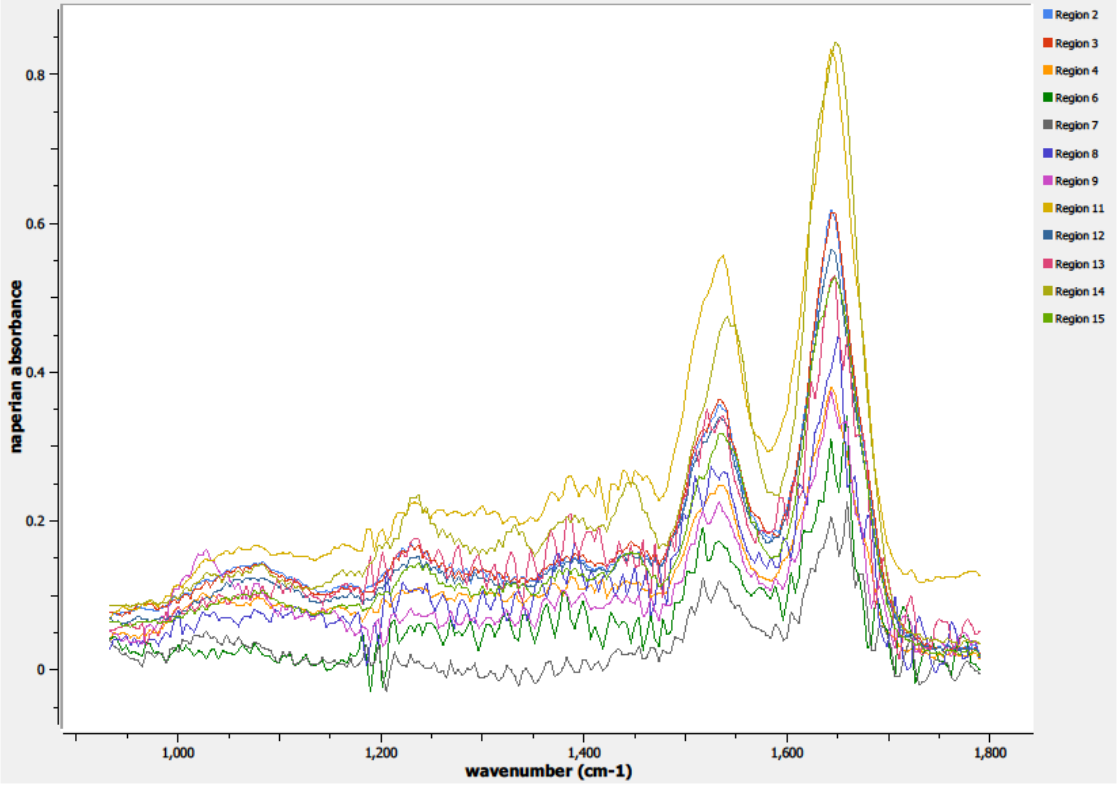



**Figure 0.11. Point N Spectra**

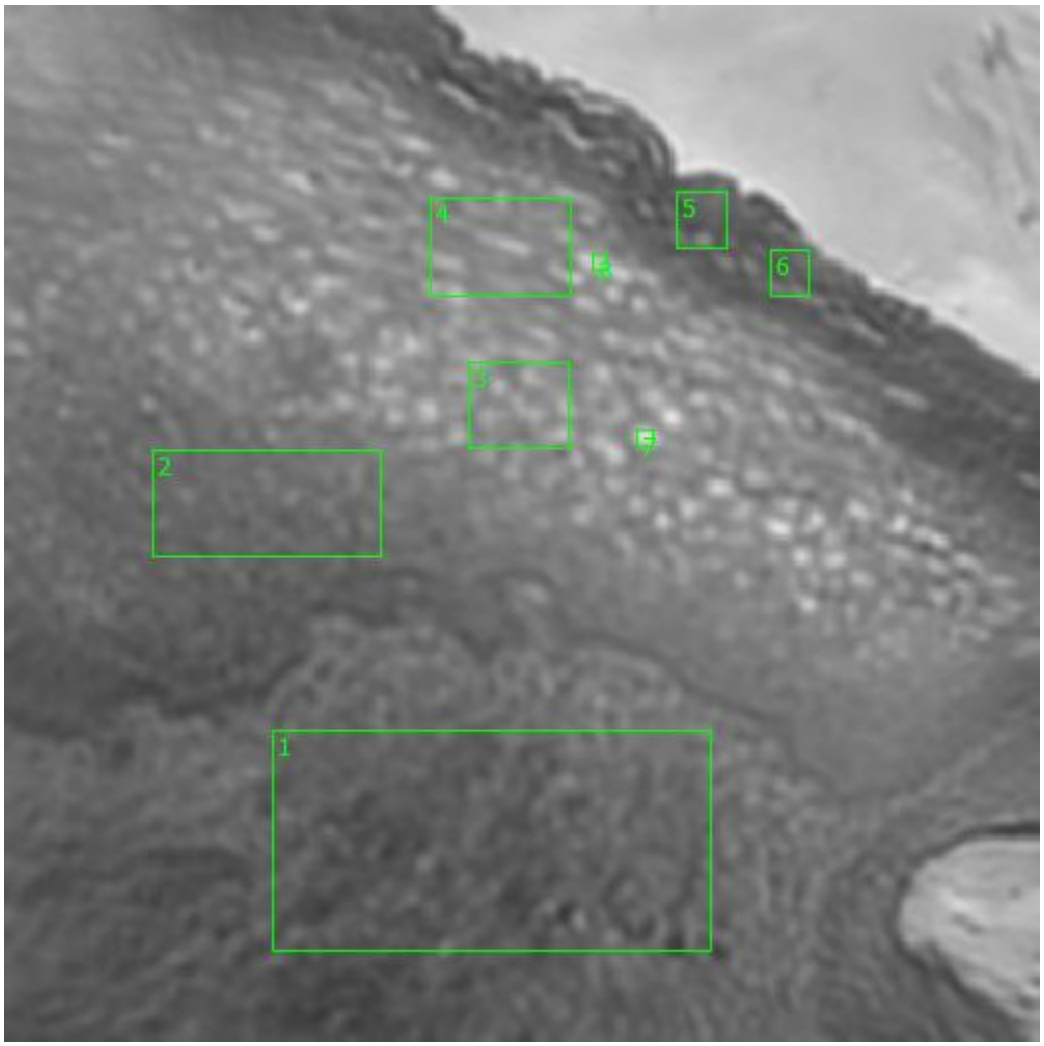



**Figure 0.12. Point O Regions of Interest**

**Region 1: basement membrane**
**Region 2: parabasal layer**
**Region 3: columnar epithelium**
**Region 4: stratified squamous epithelium**
**Regions 5, 6: stratified squamous epithelium, epithelial surface**
**Region 7: individual columnar epithelial cell**
**Region 8: individual squamous epithelial cell**

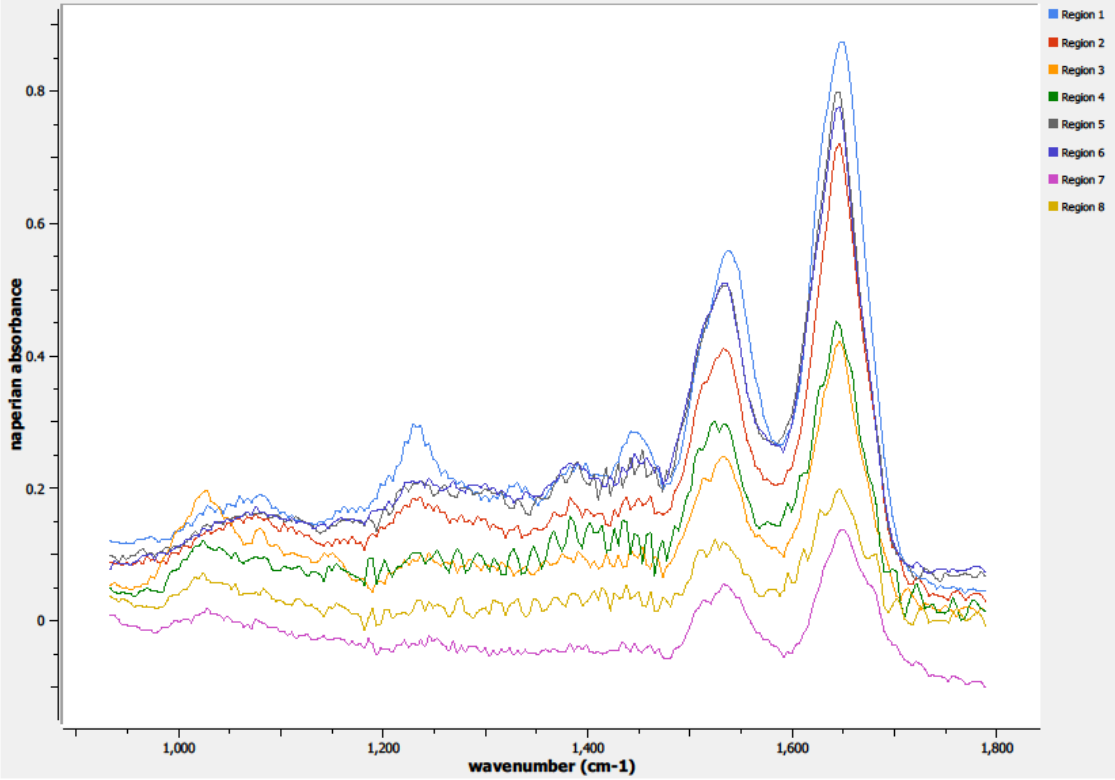



**Figure 0.13. Point O Spectra**

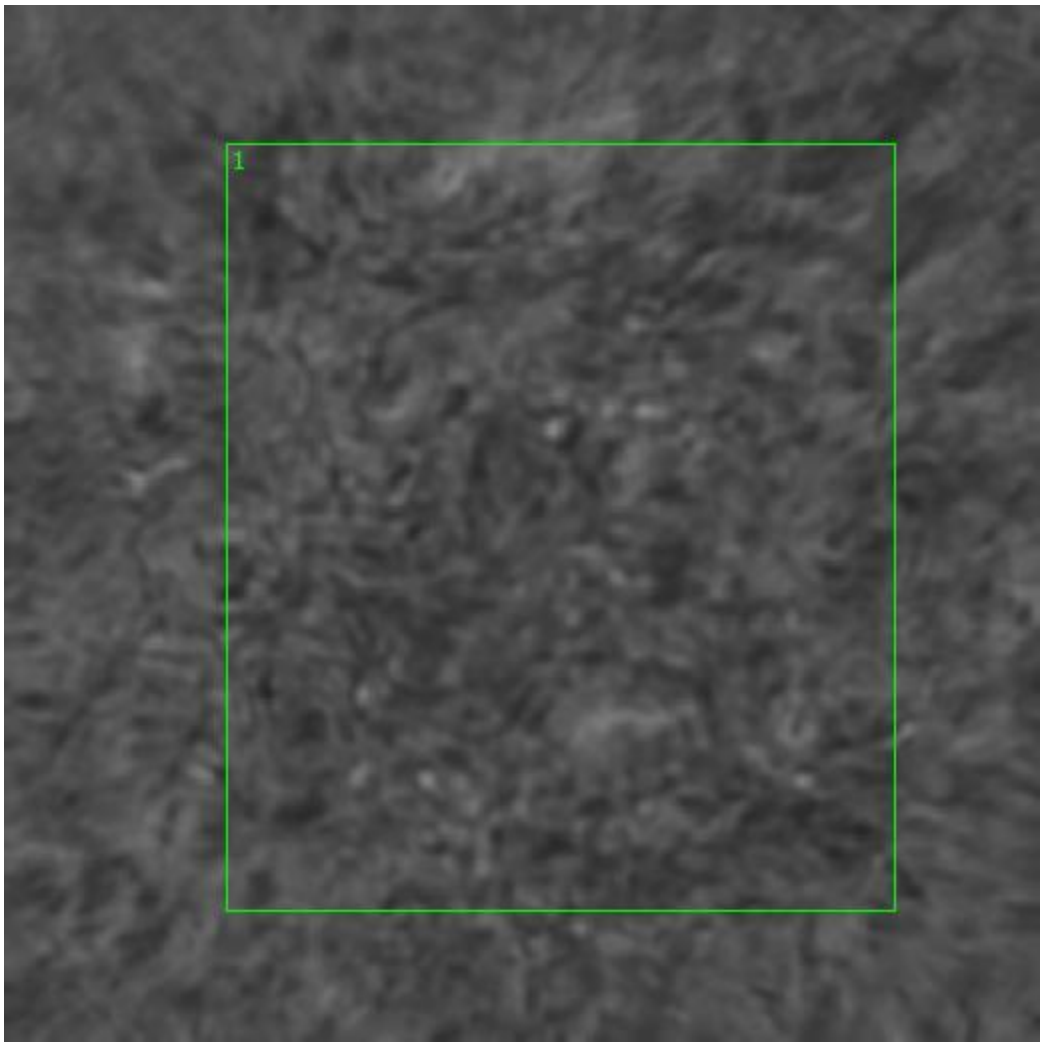



**Figure 0.14. Point P Regions of Interest**
**Region 1: lamina propia**

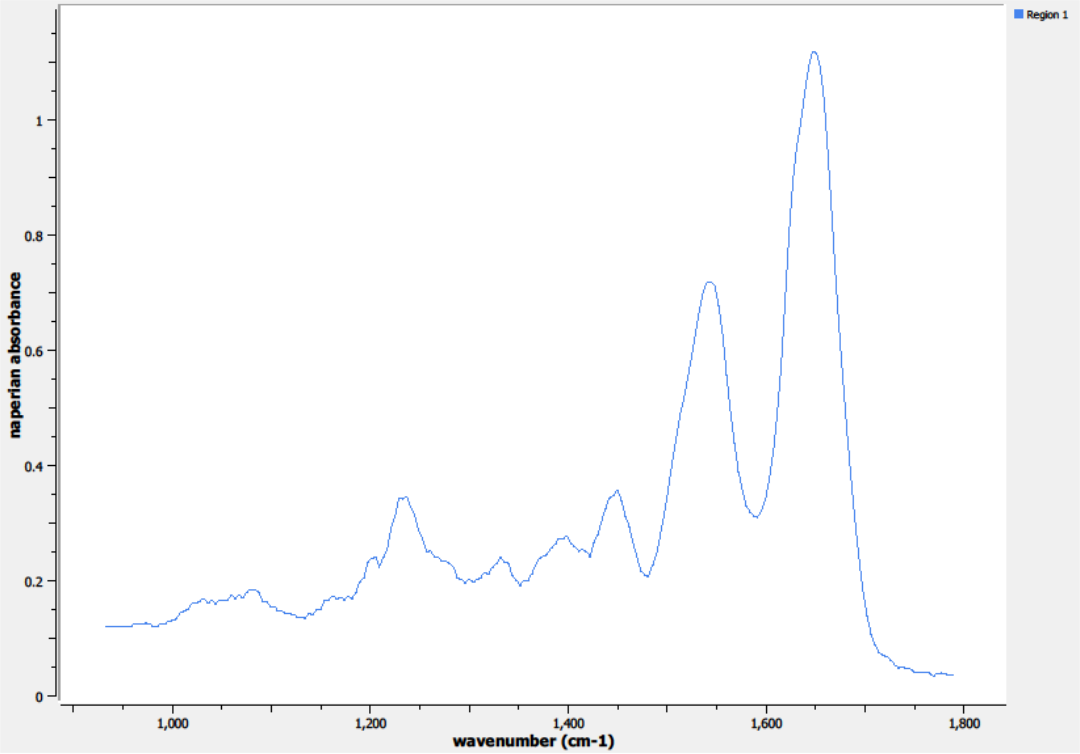



**Figure 0.15. Point P Spectra**

## A3.06-11-A257a

| | |
|---|---|
| Unique ID | 06-11-A257a |
| Age | 33 |
| Race | black |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

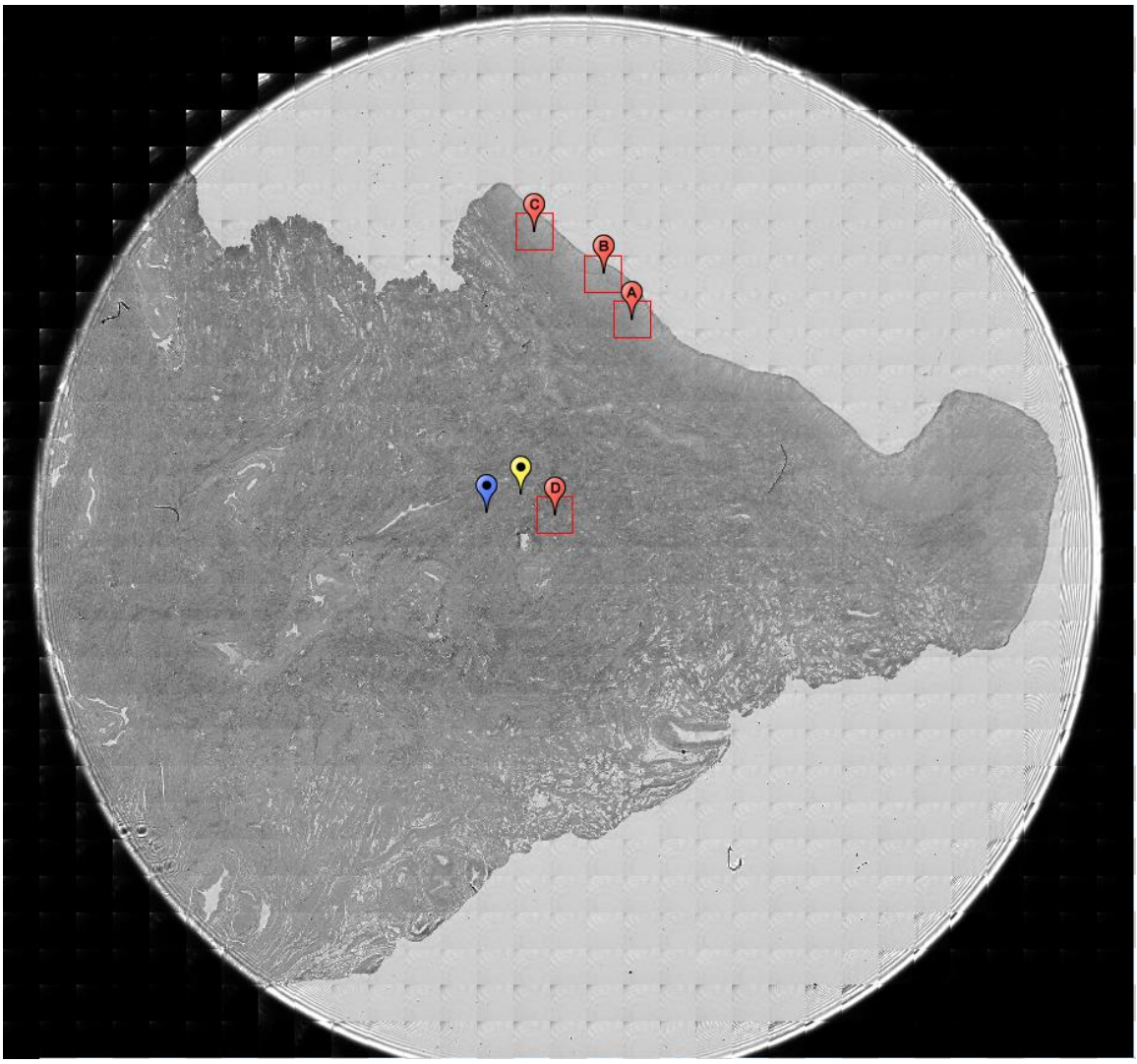



**Figure 0.16. Specimen 06-11-A257a, normal cervical tissue.**

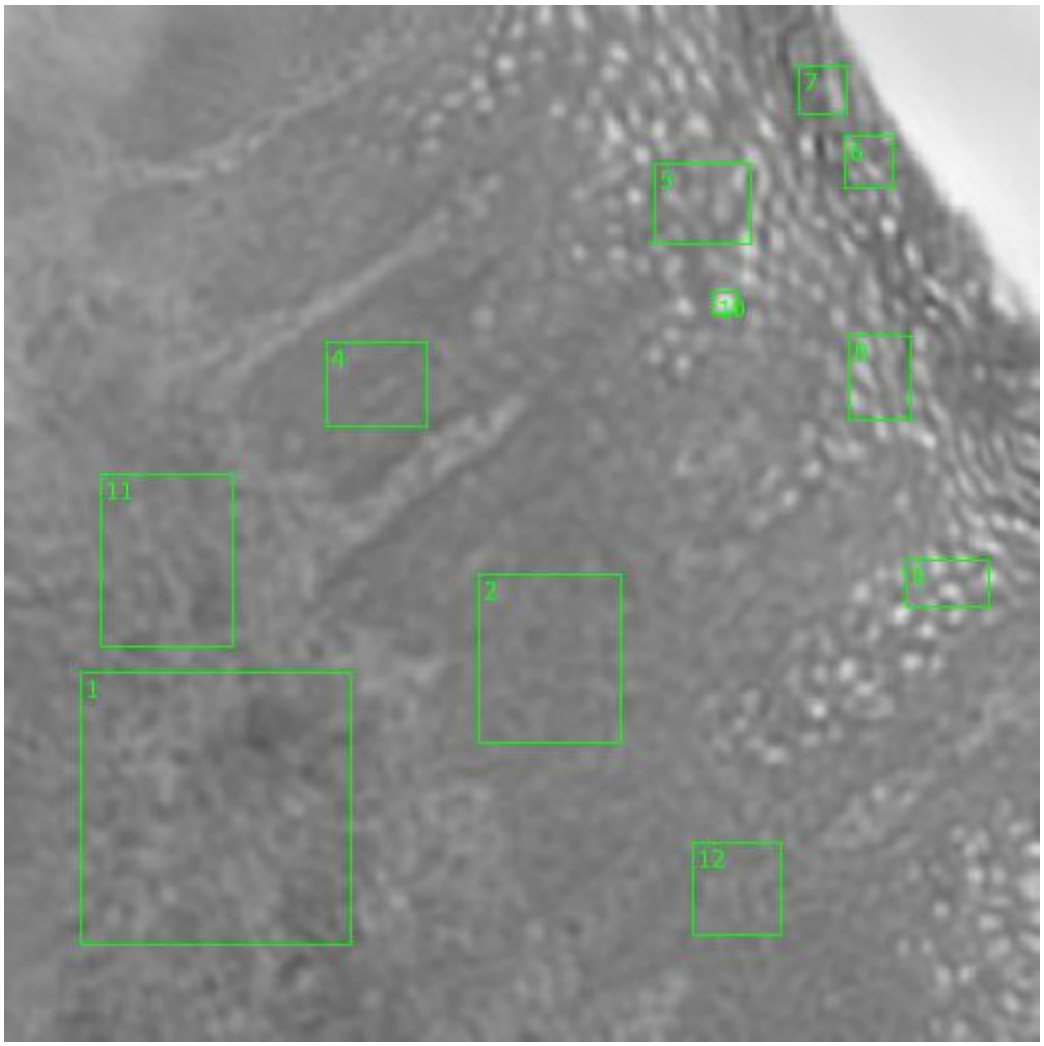



**Figure 0.17. Point A Regions of Interest**
**Regions 1, 11: basement membrane, lamina propia**
**Regions 2, 4, 12: basal layer, parabasal layer**
**Regions 5, 9: columnar epithelium**
**Regions 6, 7: stratified squamous epithelium, epithelial surface**
**Region 8: statified squamous epithelium**
**Region 10: individual columnar epithelial cell**

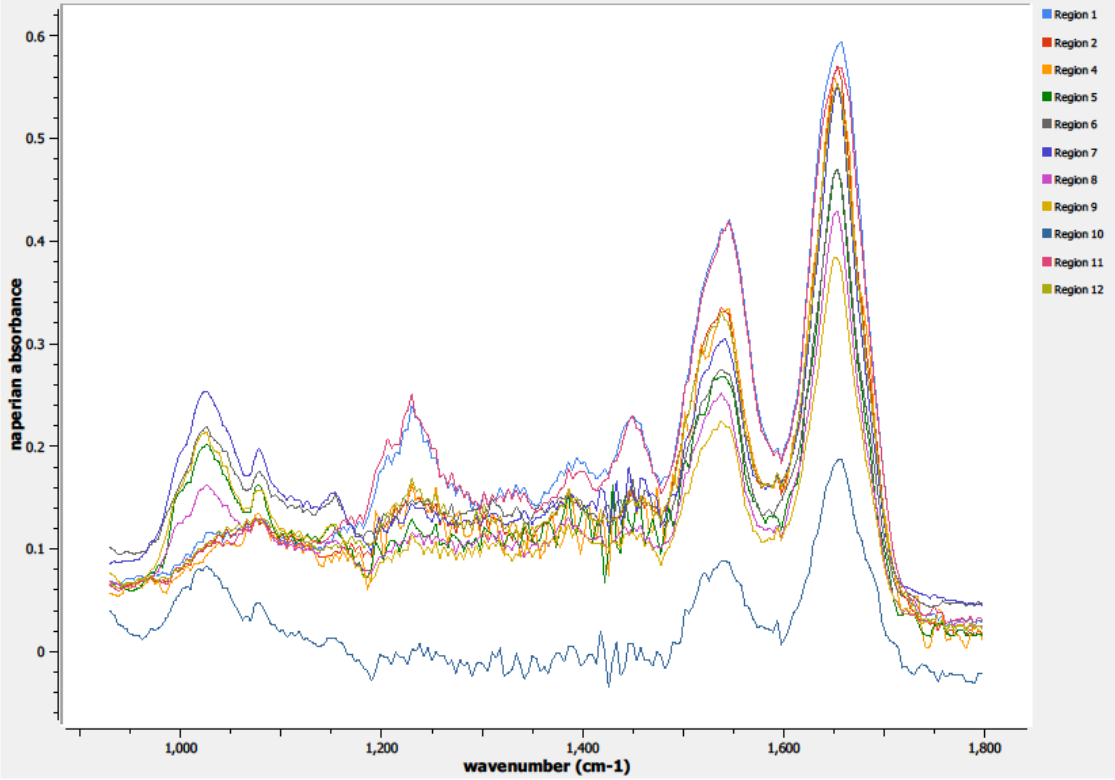



**Figure 0.18. Point A Spectra**

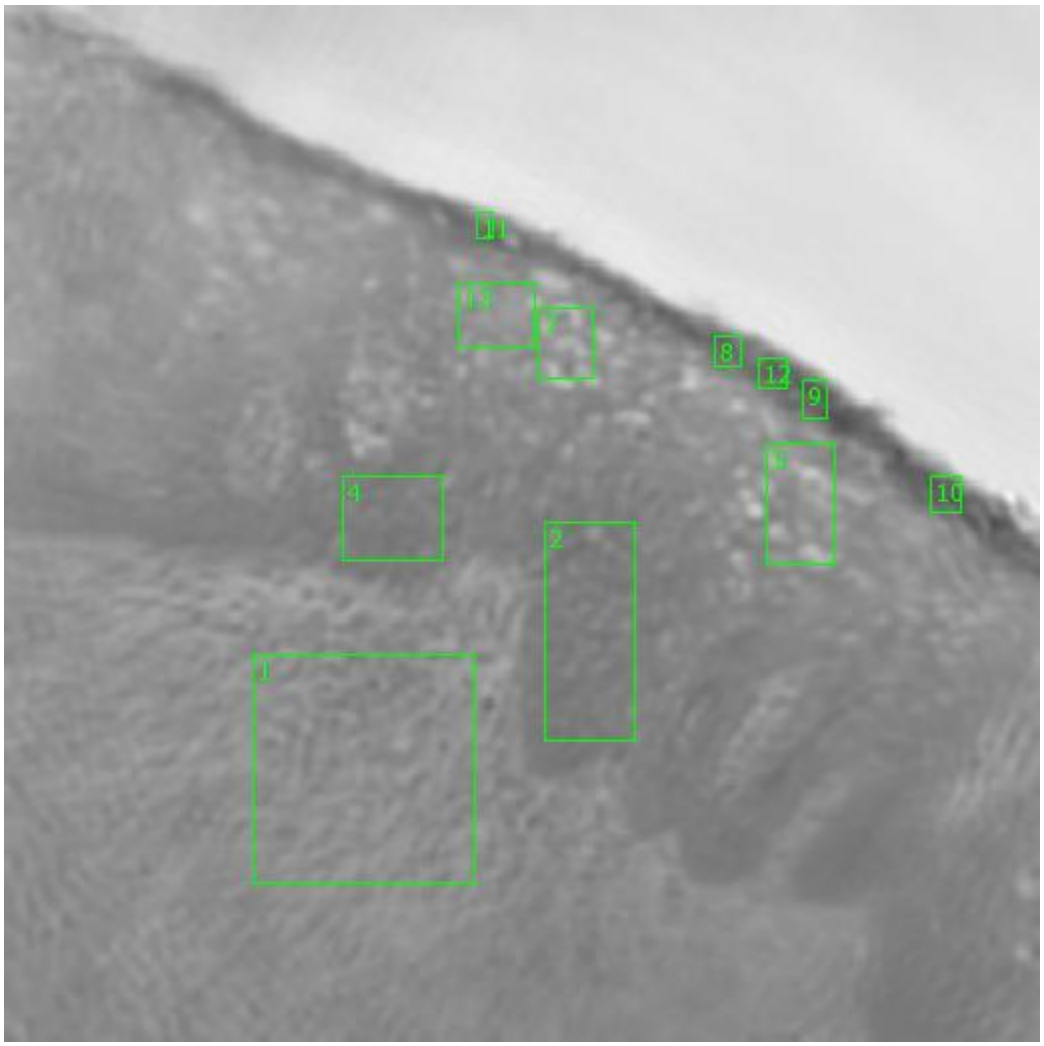



**Figure 0.19. Point B Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 4: basal layer, parabasal layer**
**Regions 5, 7, 13: columnar epithelium**
**Regions 8, 9, 10, 11, 12: stratified squamous epithelium, epithelial surface**

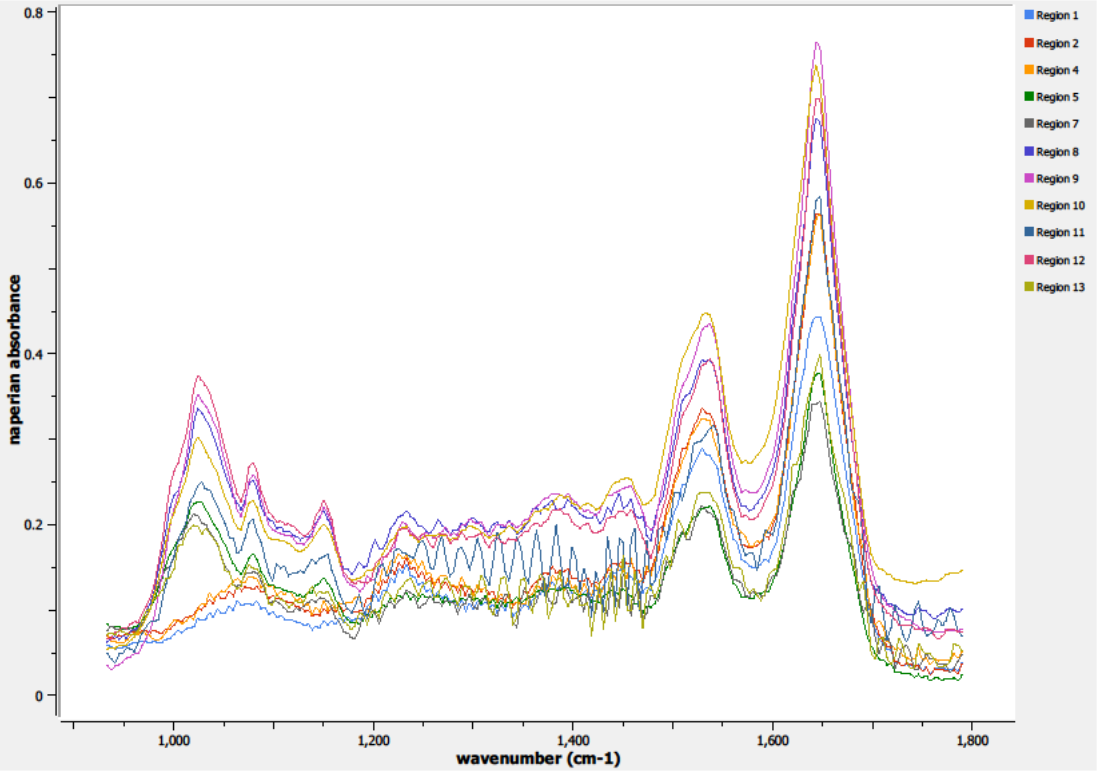



**Figure 0.20. Point B Spectra**

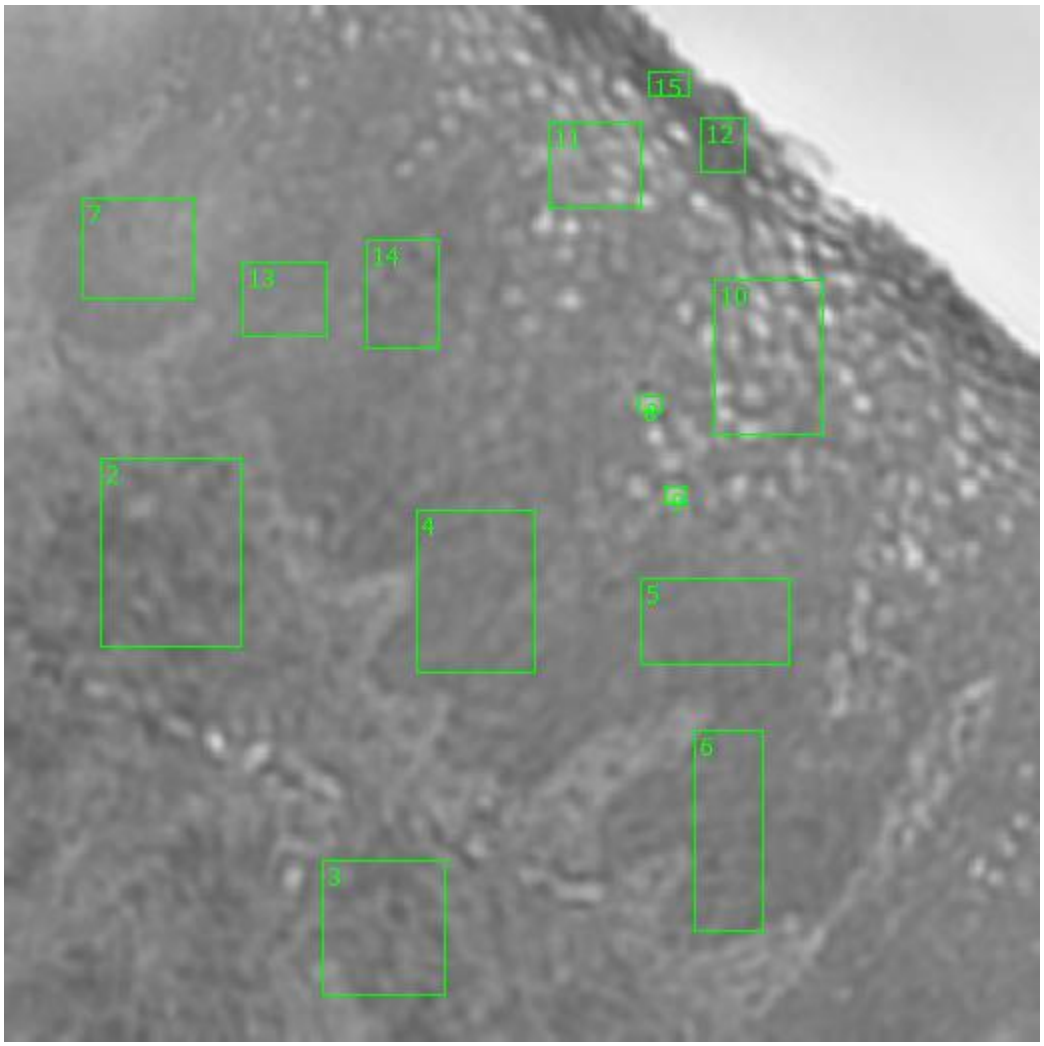



**Figure 0.21. Point C Regions of Interest**
**Regions 2, 3: basement membrane, lamina propia**
**Regions 4, 5, 6, 7, 13, 14: basal layer, parabasal layer**
**Regions 8, 9: individual columnar epithelial cells**
**Region 10: columnar epithelium**
**Region 11: stratified squamous epithelium**
**Regions 12, 15: stratified squamous epithelium, epithelial surface**

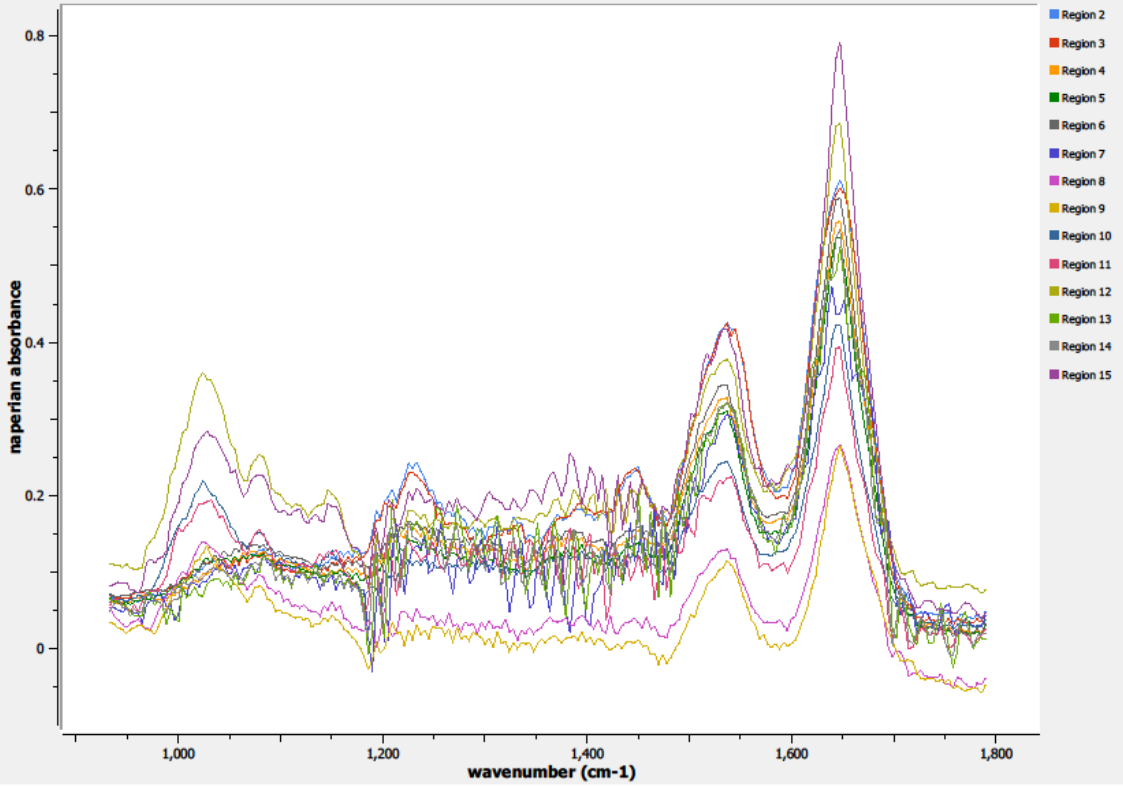



**Figure 0.22. Point C Spectra**

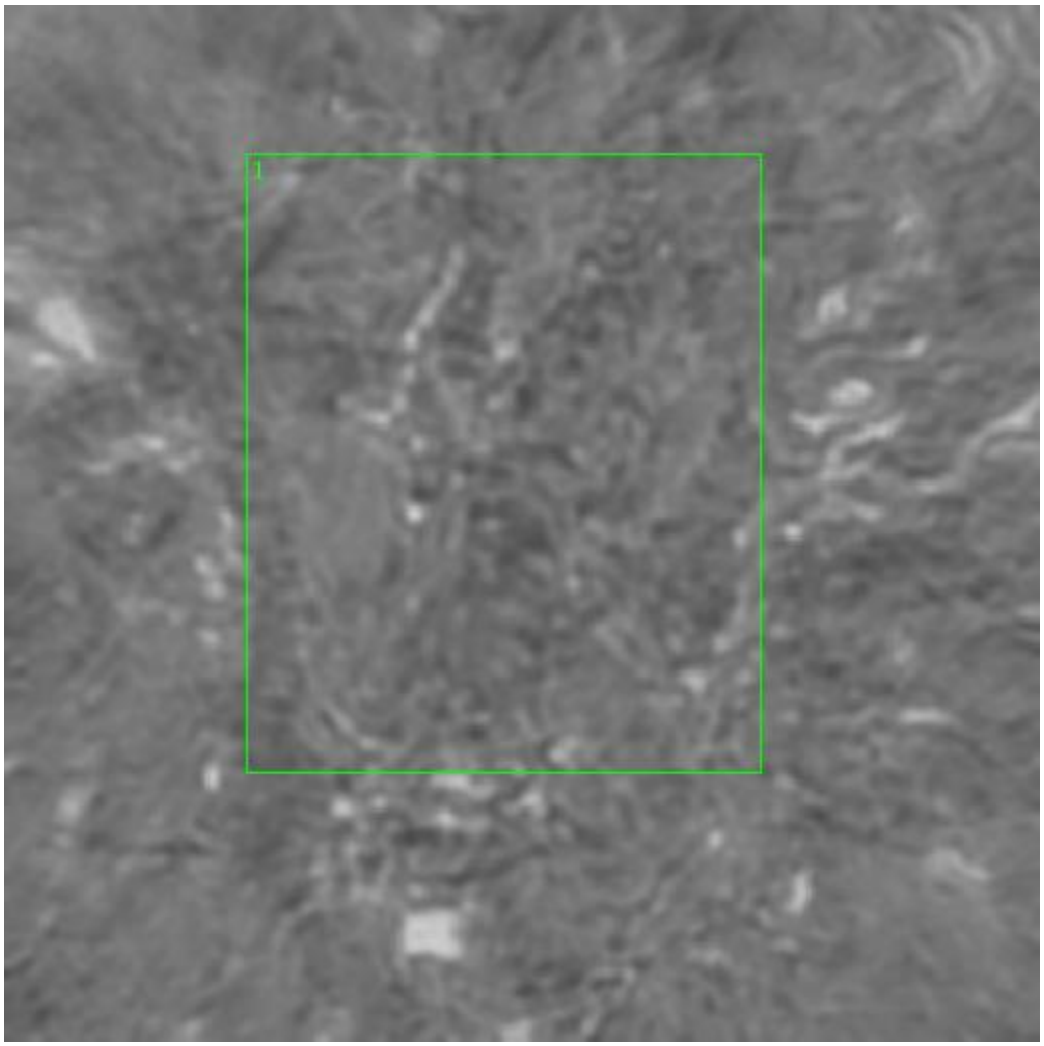



**Figure 0.23. Point D Regions of Interest**
**Region 1: lamina propia**

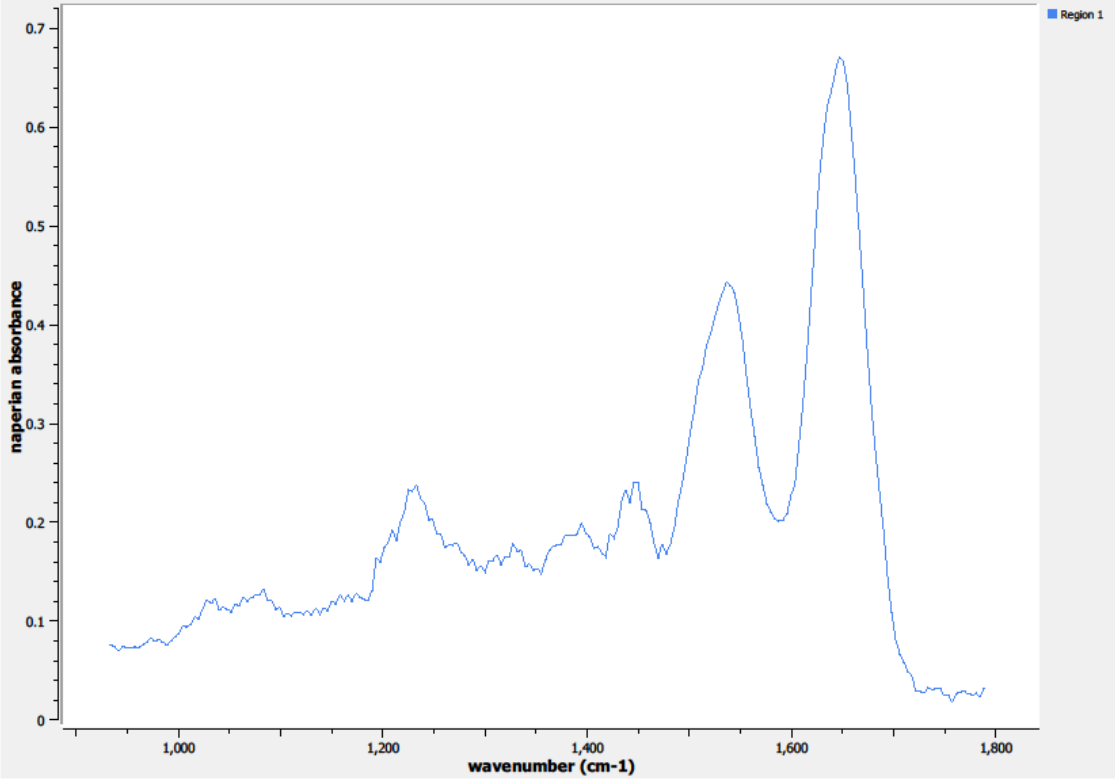



**Figure 0.24. Point D Spectra**

## A4.06-11-A378a

| | |
|---|---|
| Unique ID | 06-11-A378a |
| Age | 38 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | may not use |

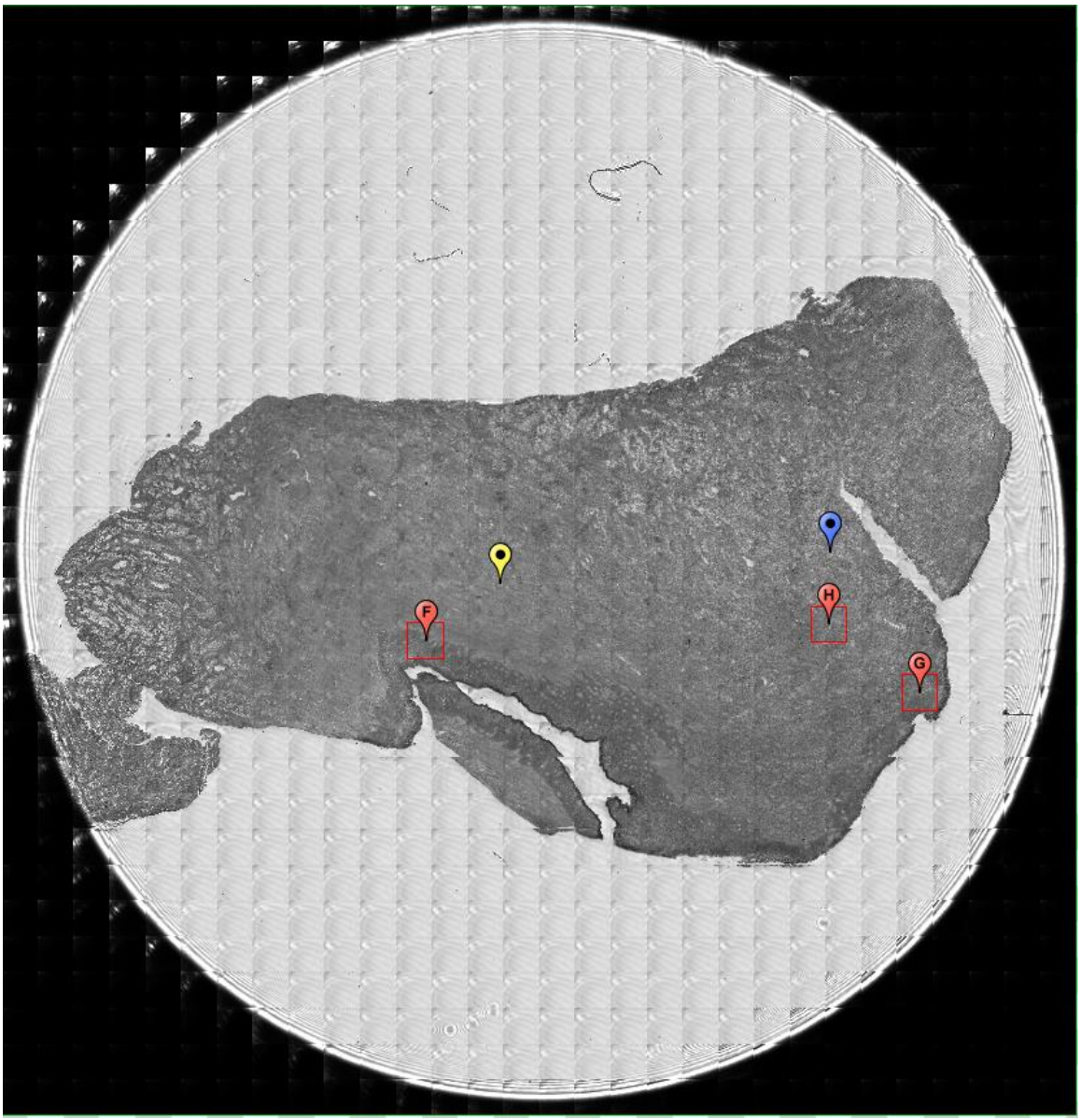



**Figure 0.25. Specimen 06-11-A378a, normal cervical tissue.**

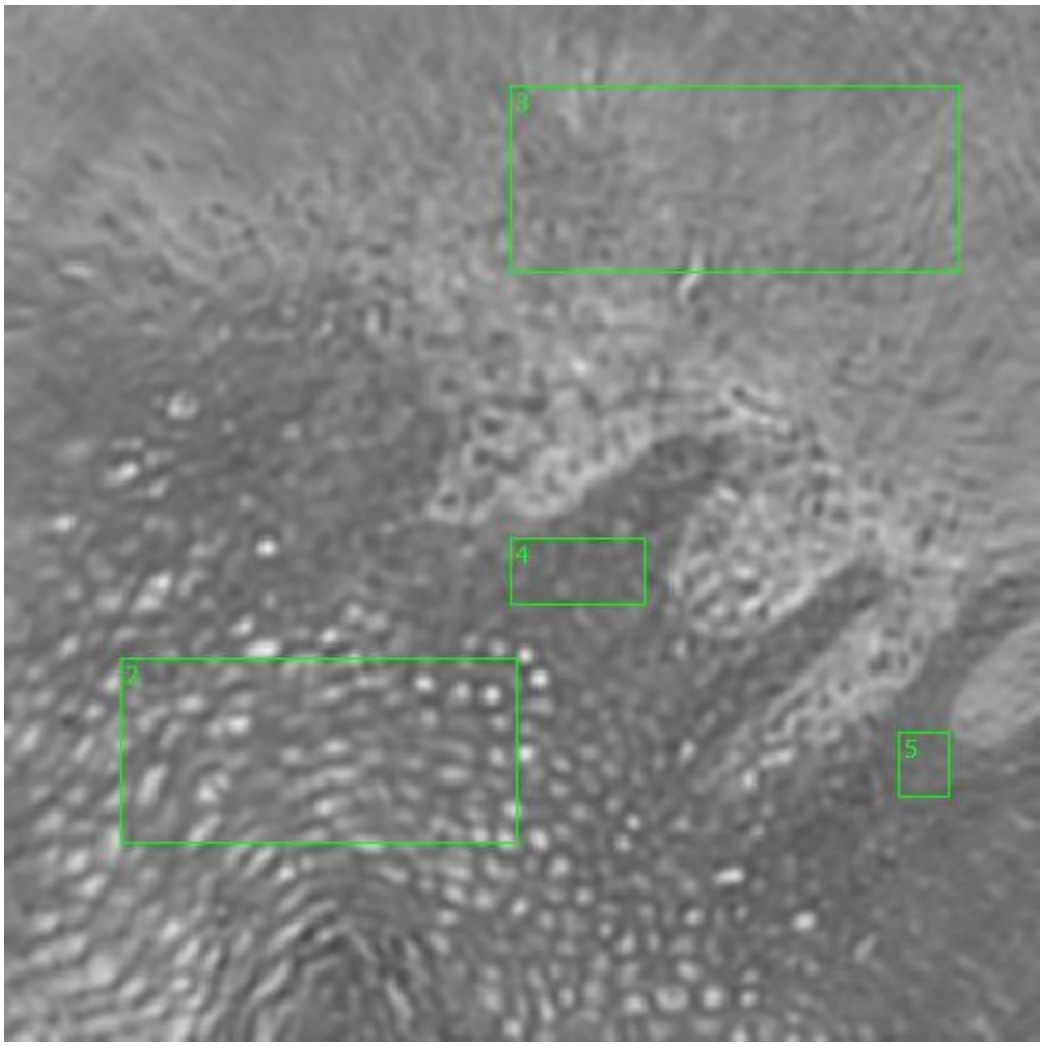



**Figure 0.26. Point F Regions of Interest**
**Region 2: statified squamous epithelium**
**Region 3: basement membrane, lamina propia**
**Regions 4, 5: basal layer**

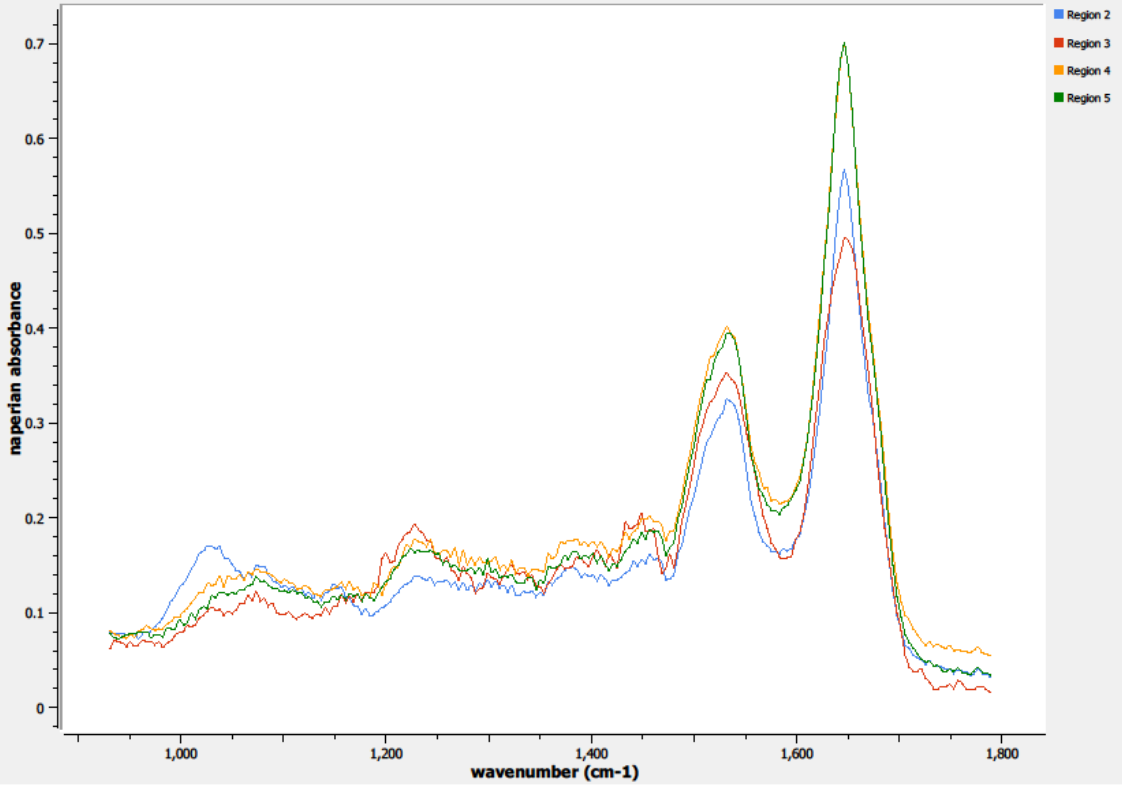



**Figure 0.27. Point F Spectra**

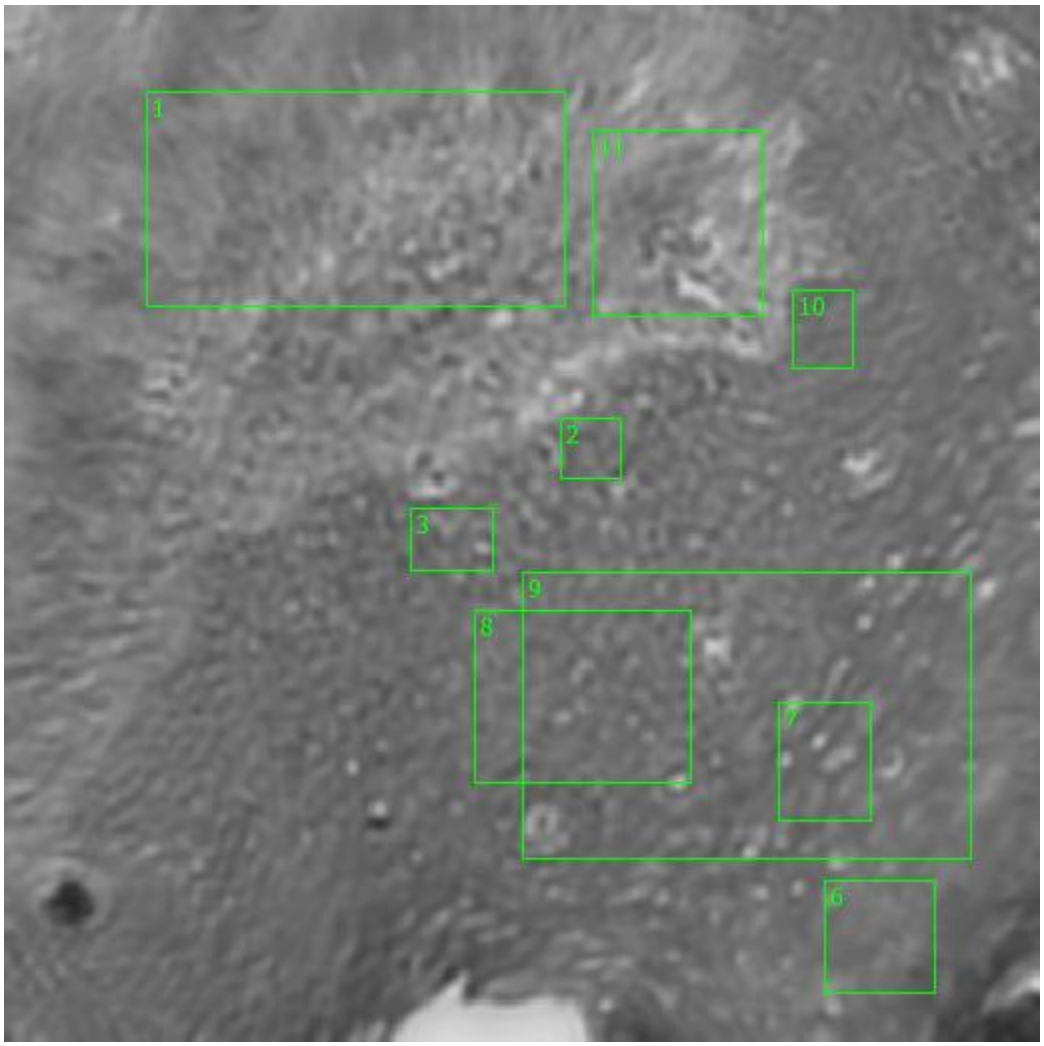



**Figure 0.28. Point G Regions of Interest**
**Regions 1, 11: basement membrane, lamina propia**
**Regions 2, 3, 10: basal layer, parabasal layer**
**Regions 6, 7: statified squamous epithelium**
**Region 8: columnar epithelium**
**Region 9: columnar epithelium, statified squamous epithelium**

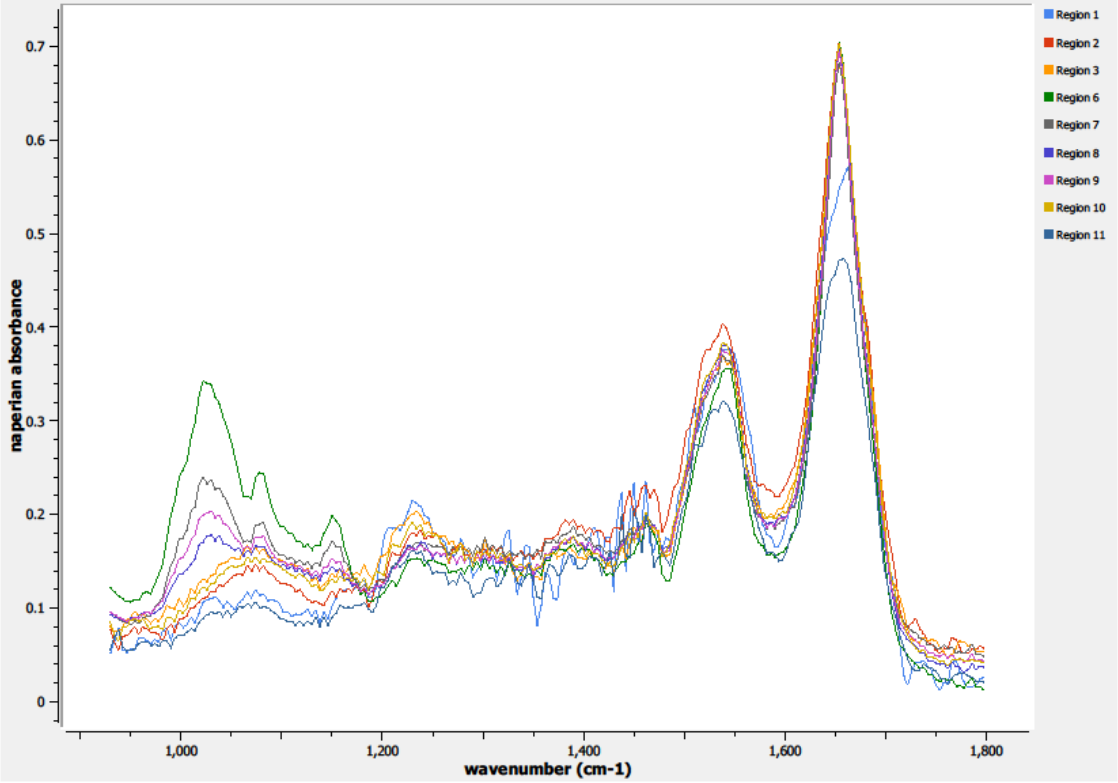



**Figure 0.29. Point G Spectra**

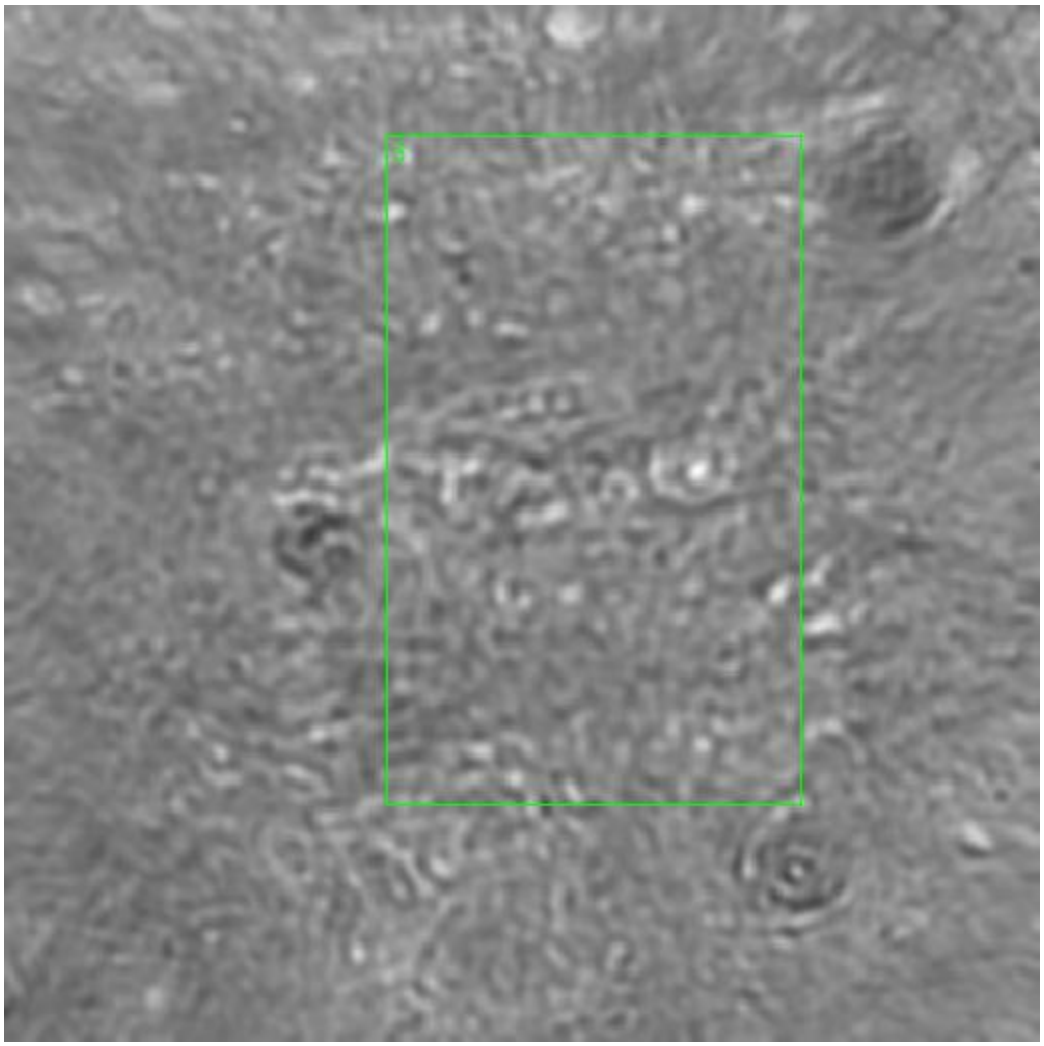


**Figure 0.30. Point H Regions of Interest**
**Regions 3: lamina propia**

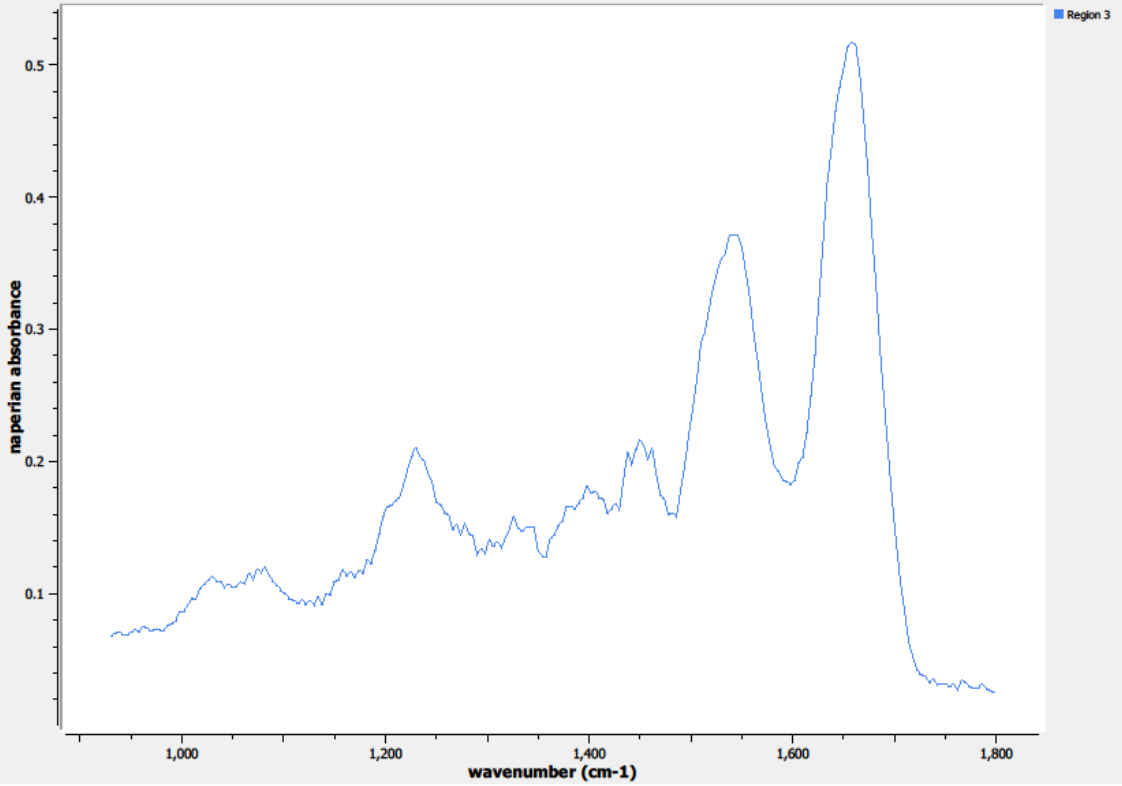



**Figure 0.31. Point H Spectra**

## A5.06-11-A475a

| | |
|---|---|
| Unique ID | 06-11-A475a |
| Age | 31 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | none |

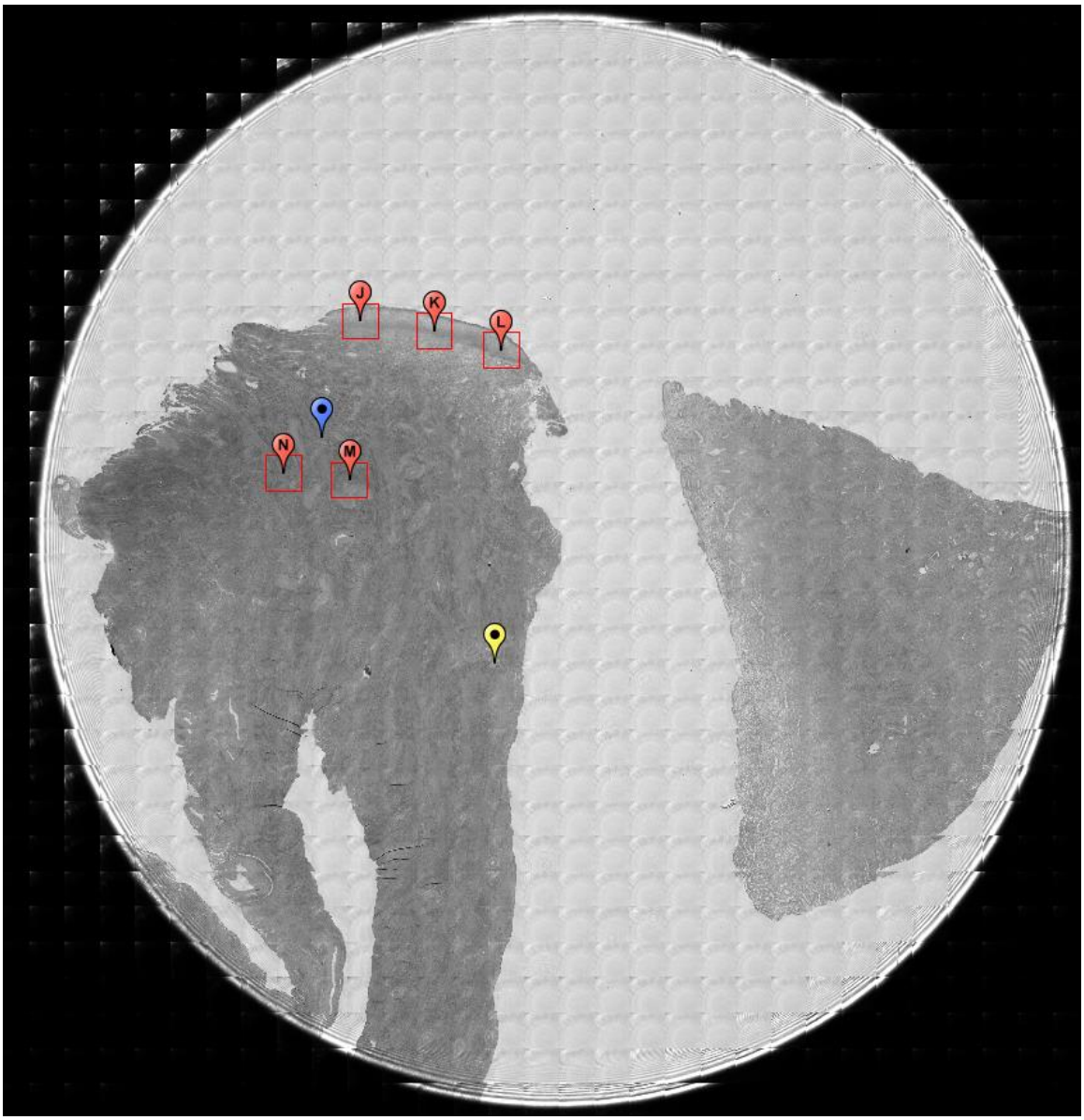



**Figure 0.32. Specimen 06-11-A475a, normal cervical tissue.**

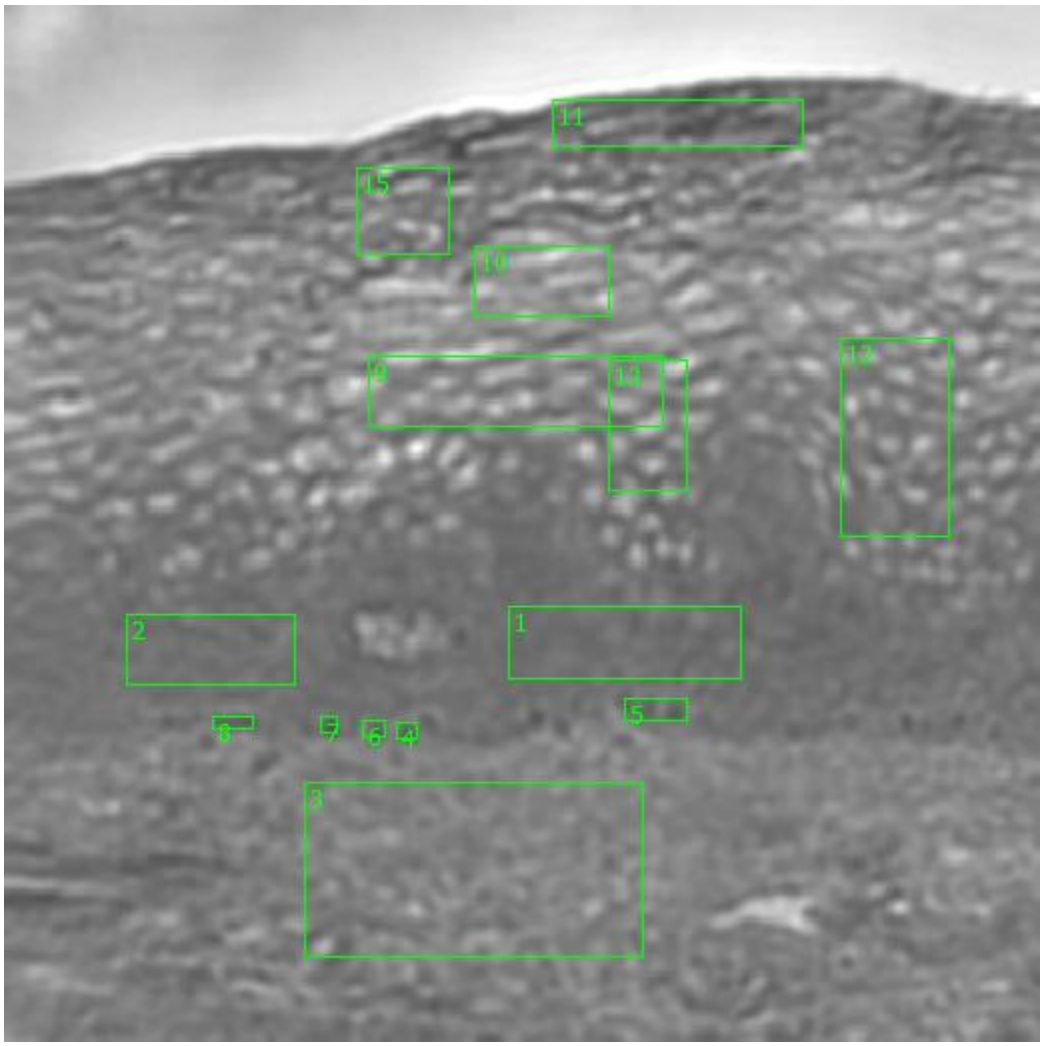



**Figure 0.33. Point J Regions of Interest**
**Regions 1, 2: parabasal layer**
**Region 3: basement membrane, lamina propia**
**Regions 4, 5, 6, 7, 8: basal layer**
**Regions 9, 12, 13: columnar epithelium**
**Regions 10, 15: statified squamous epithelium**
**Region 11: stratified squamous epithelium, epithelial surface**

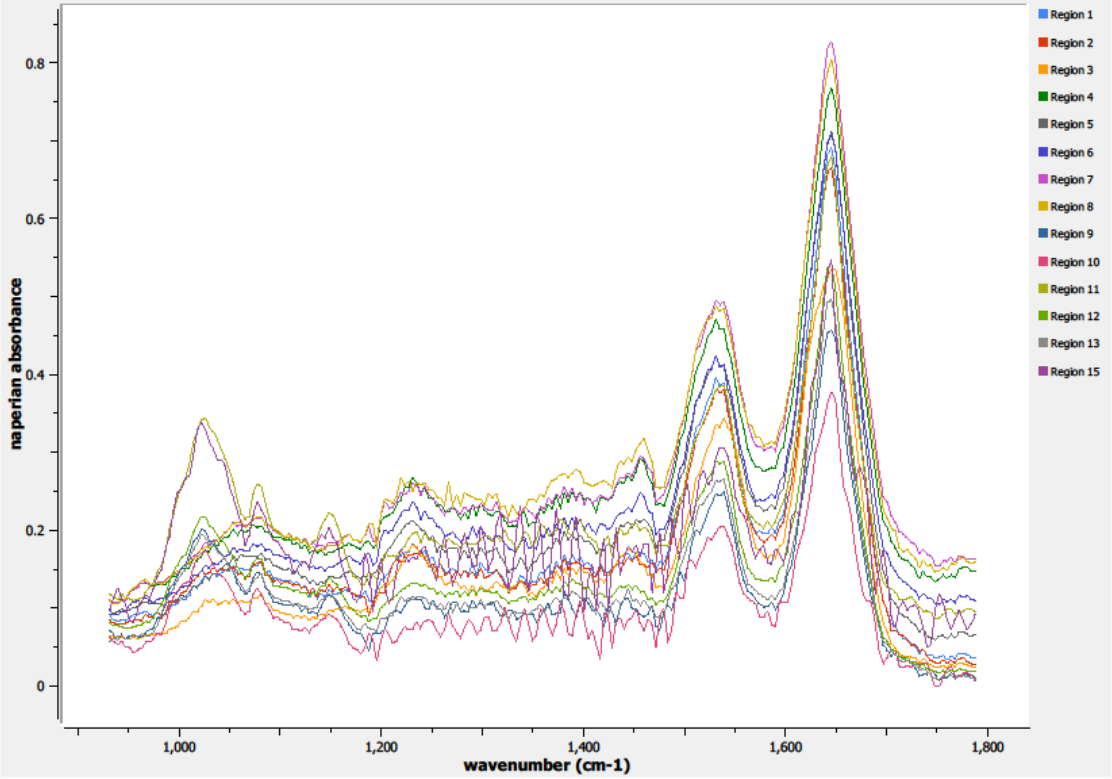



**Figure 0.34. Point J Spectra**

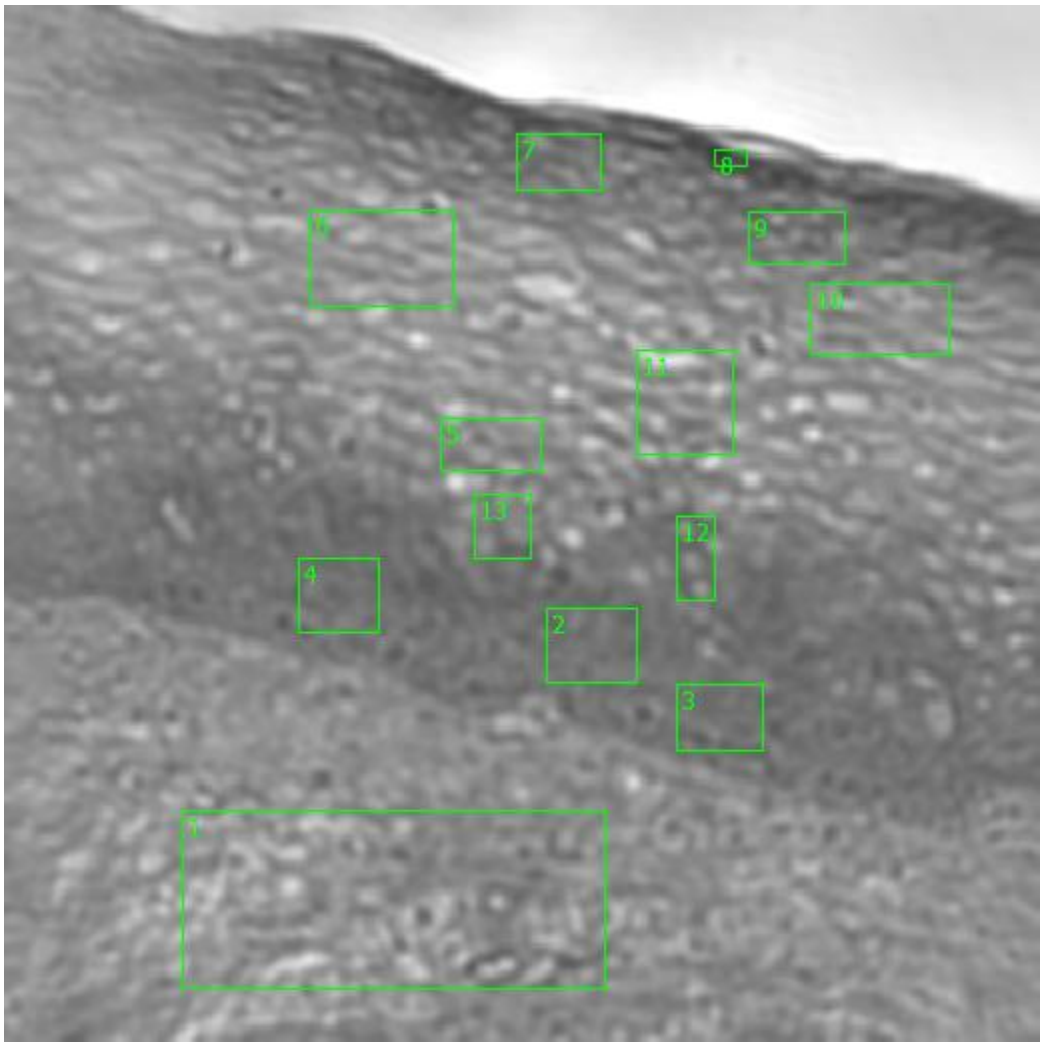



**Figure 0.35. Point K Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 3, 4: parabasal layer**
**Regions 5, 11, 12, 13: columnar epithelium**
**Regions 6, 7, 9, 10: statified squamous epithelium**
**Region 8: stratified squamous epithelium, epithelial surface**

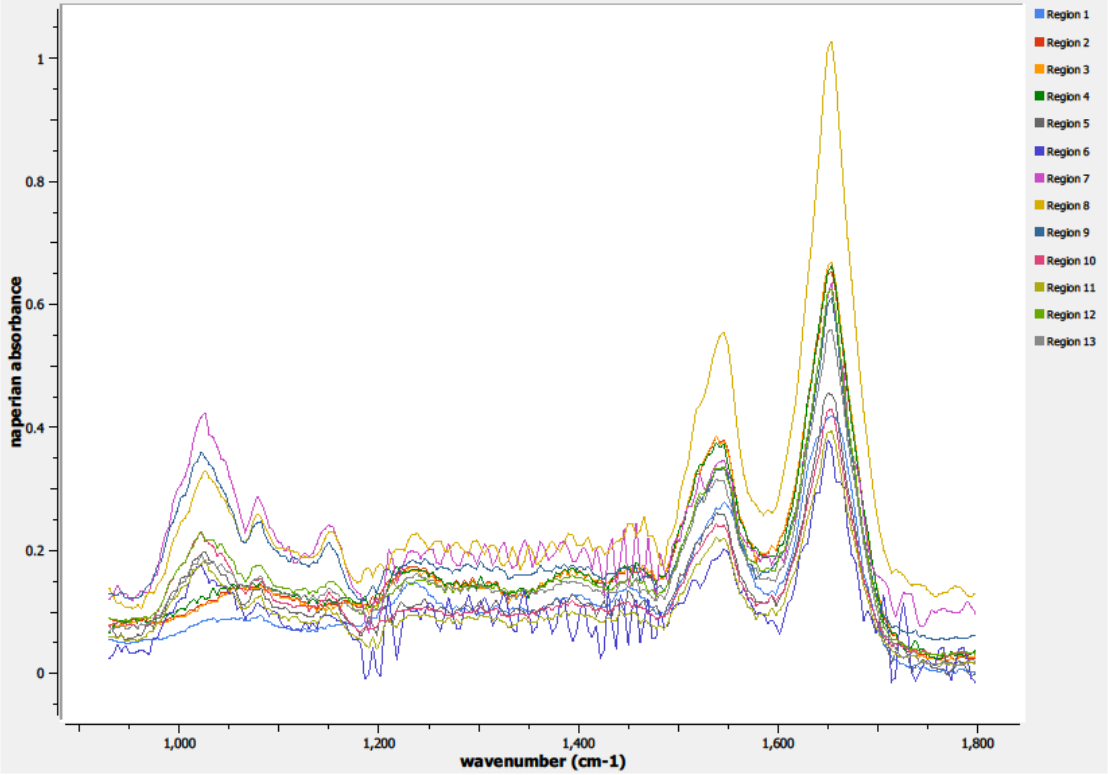



**Figure 0.36. Point K Spectra**

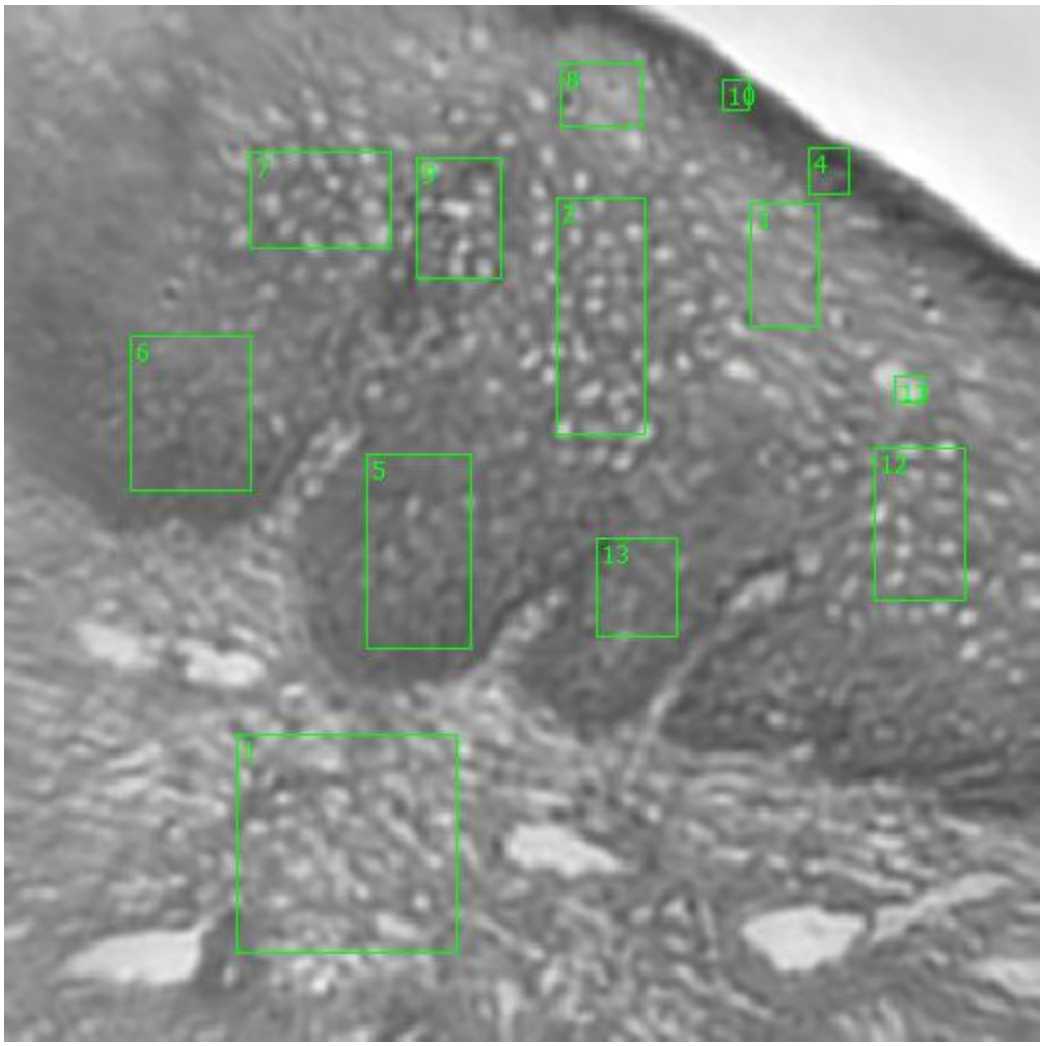



**Figure 0.37. Point L Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 7, 9, 12: columnar epithelium**
**Regions 3, 8: statified squamous epithelium**
**Regions 4, 10: stratified squamous epithelium, epithelial surface**
**Regions 5, 6, 13: parabasal layer**
**Region 11: individual columnar epithelial cell**

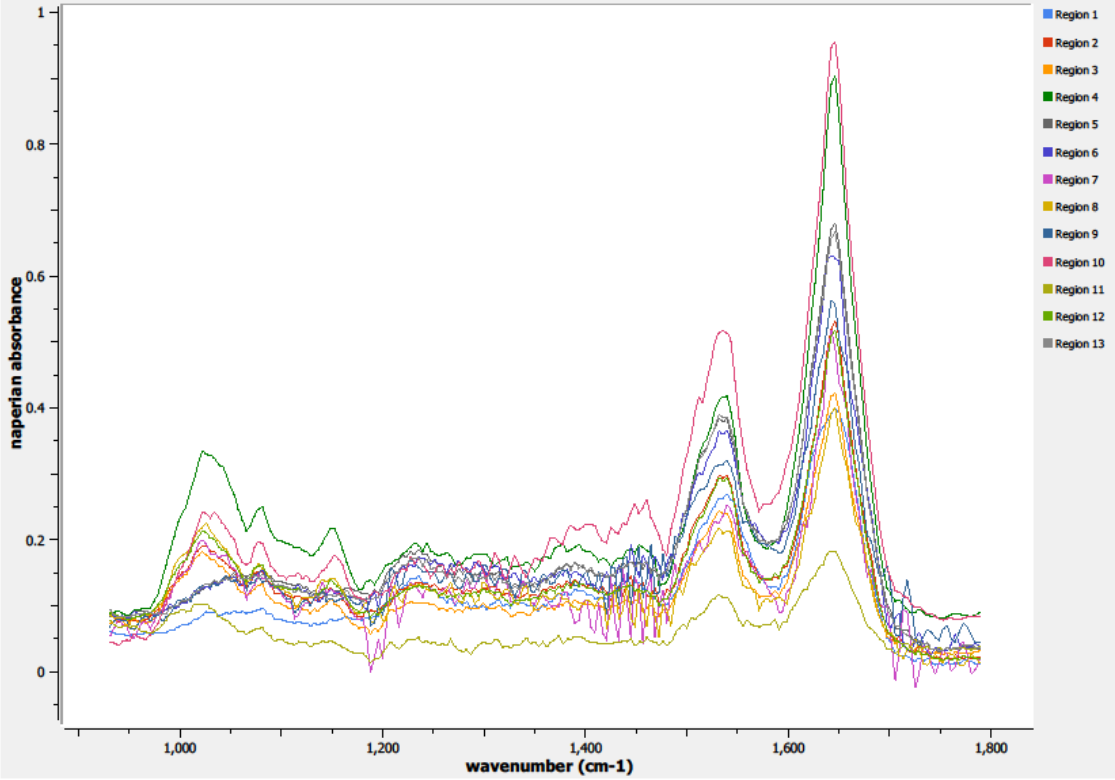



**Figure 0.38. Point L Spectra**

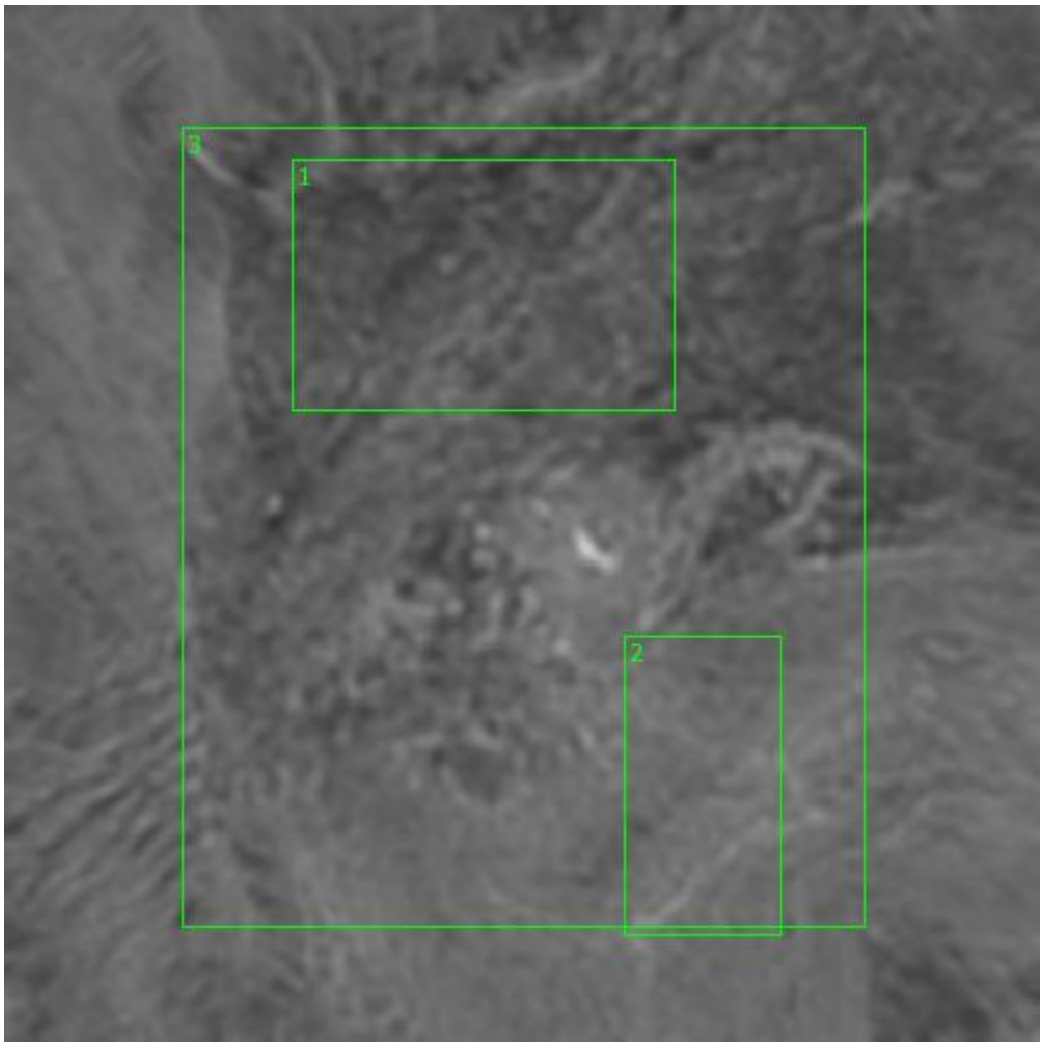



**Figure 0.39. Point M Regions of Interest**
**Regions 1, 2, 3: basement membrane, lamina propia**

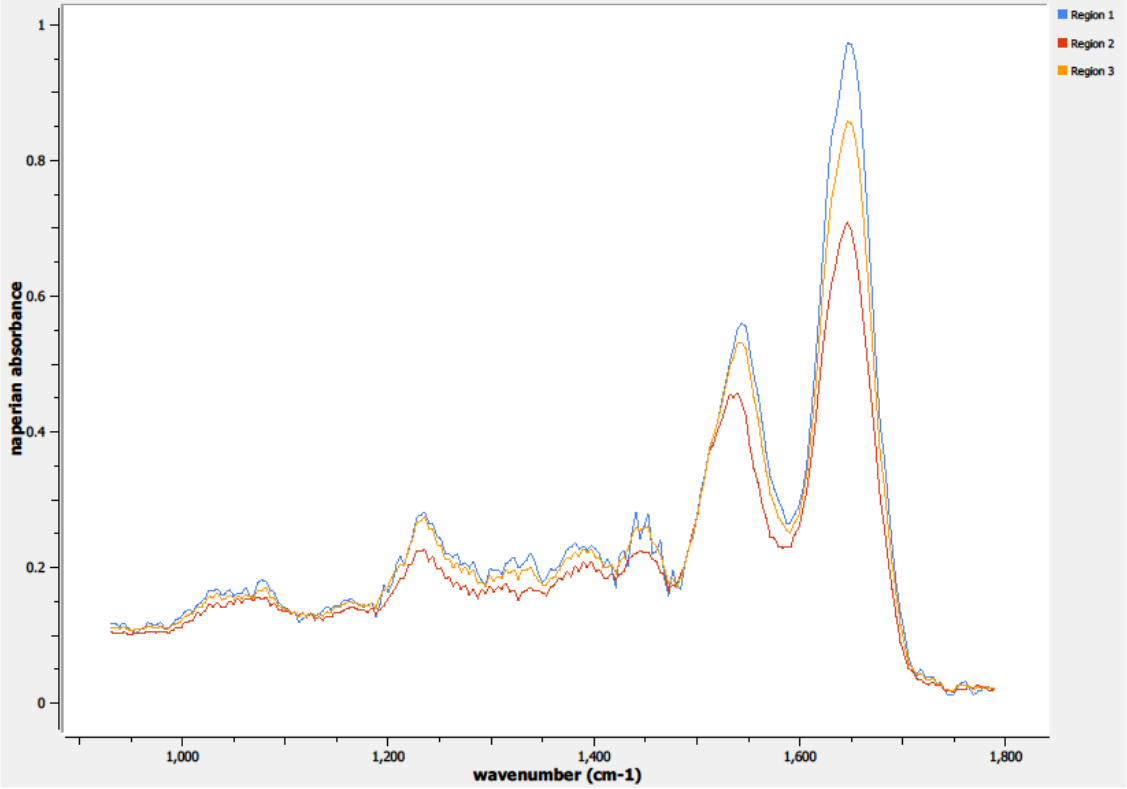



**Figure 0.40. Point M Spectra**

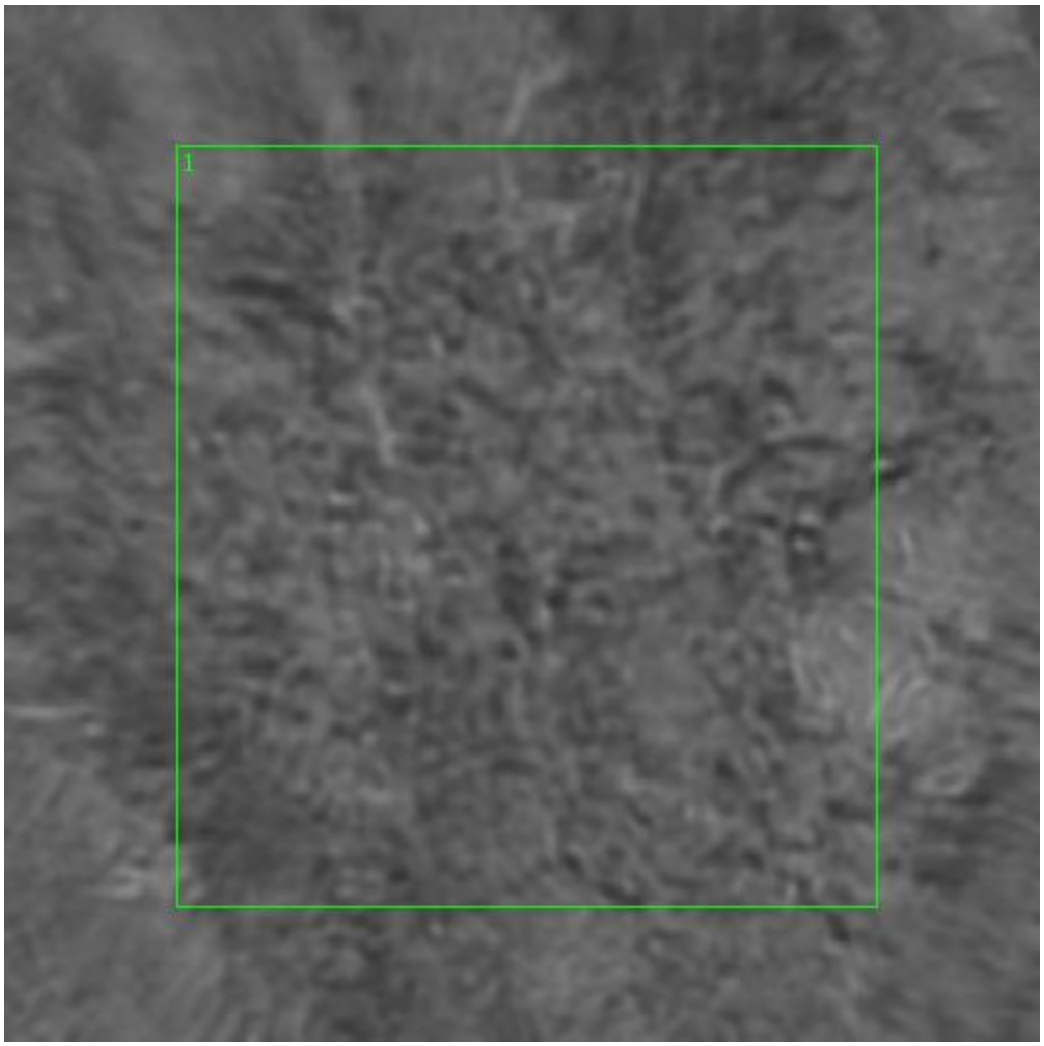



**Figure 0.41. Point N Region of Interest**
**Region 1: basement membrane, lamina propia**

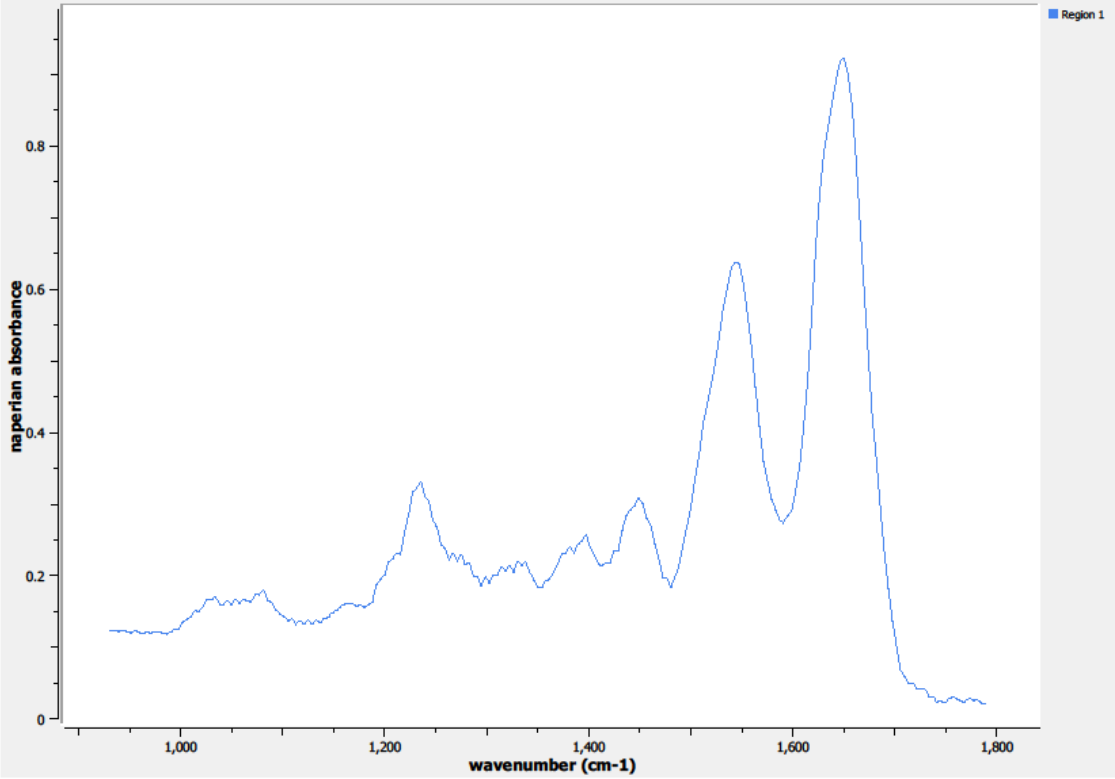



**Figure 0.42. Point N Spectra**

## A6.12-05-A041a

| | |
|---|---|
| Unique ID | 12-05-A041a |
| Age | 65 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | no clearly-defined epithelium layer |

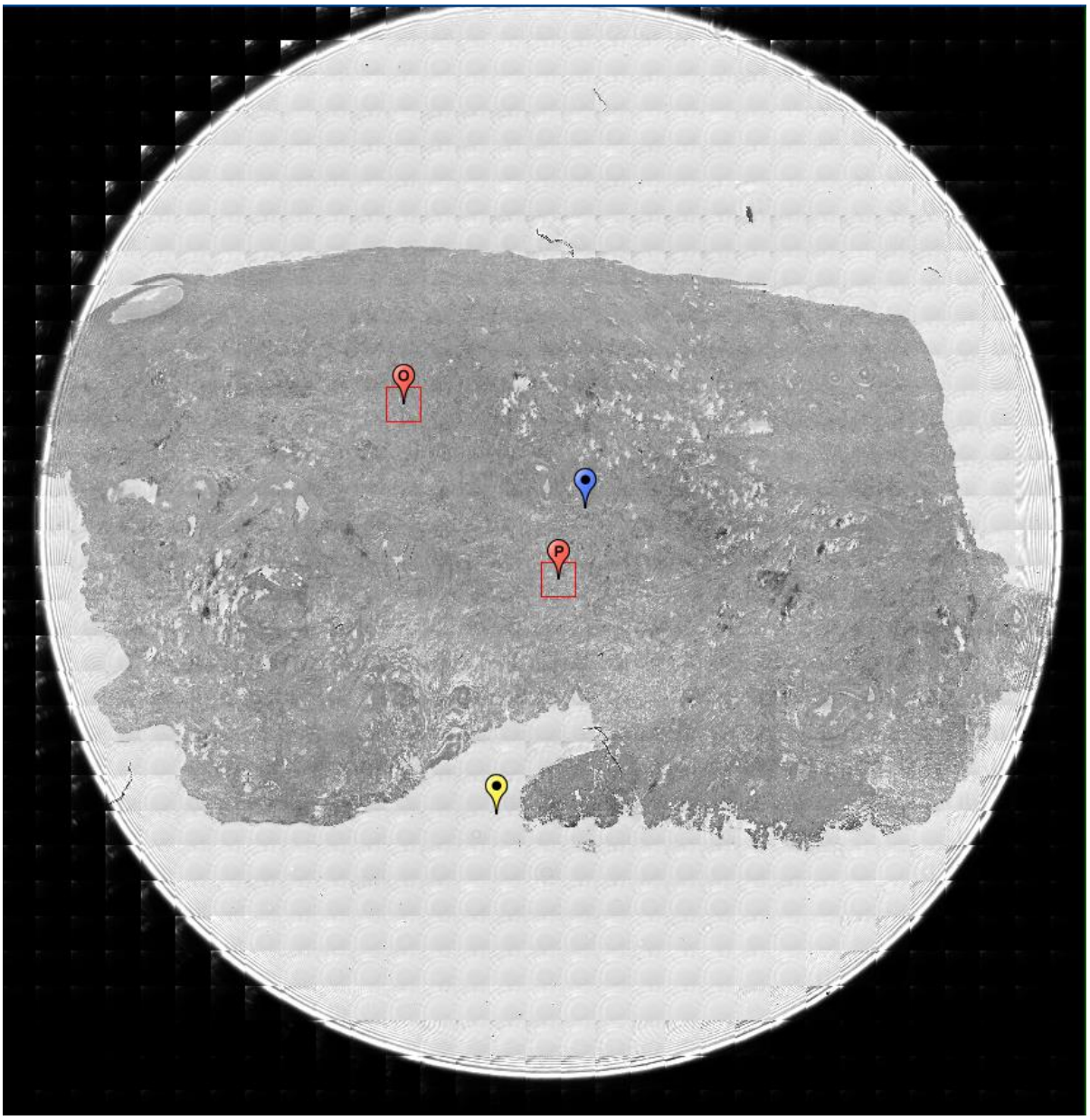



**Figure 0.43. Specimen 12-05-A041a, normal cervical tissue.**

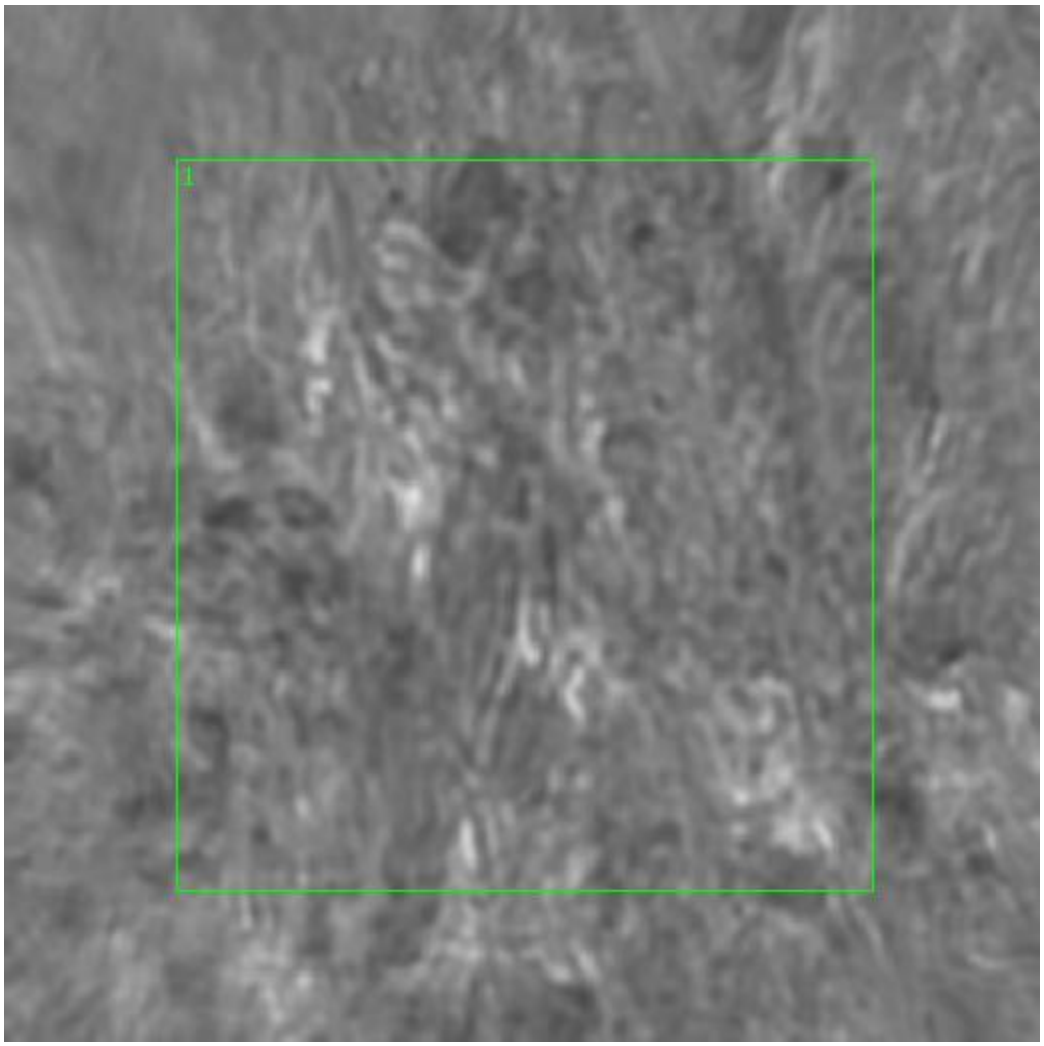



**Figure 0.44. Point O Region of Interest**
**Region 1: basement membrane, lamina propia**

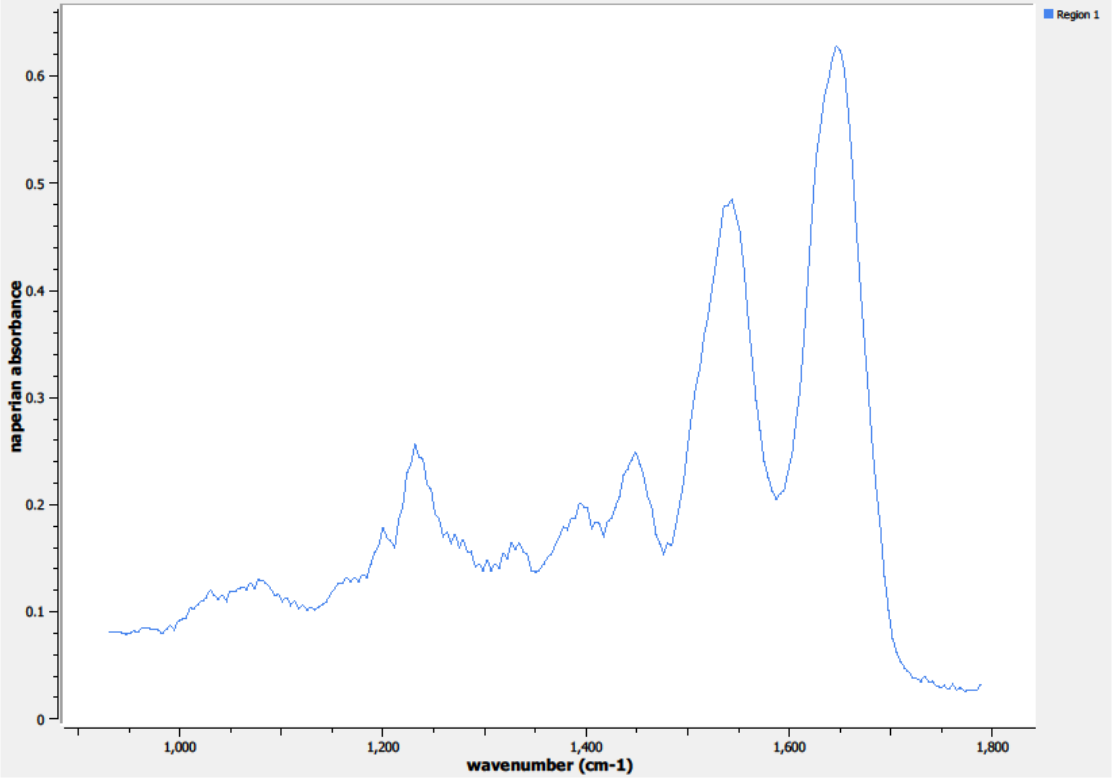



**Figure 0.45. Point O Spectra**

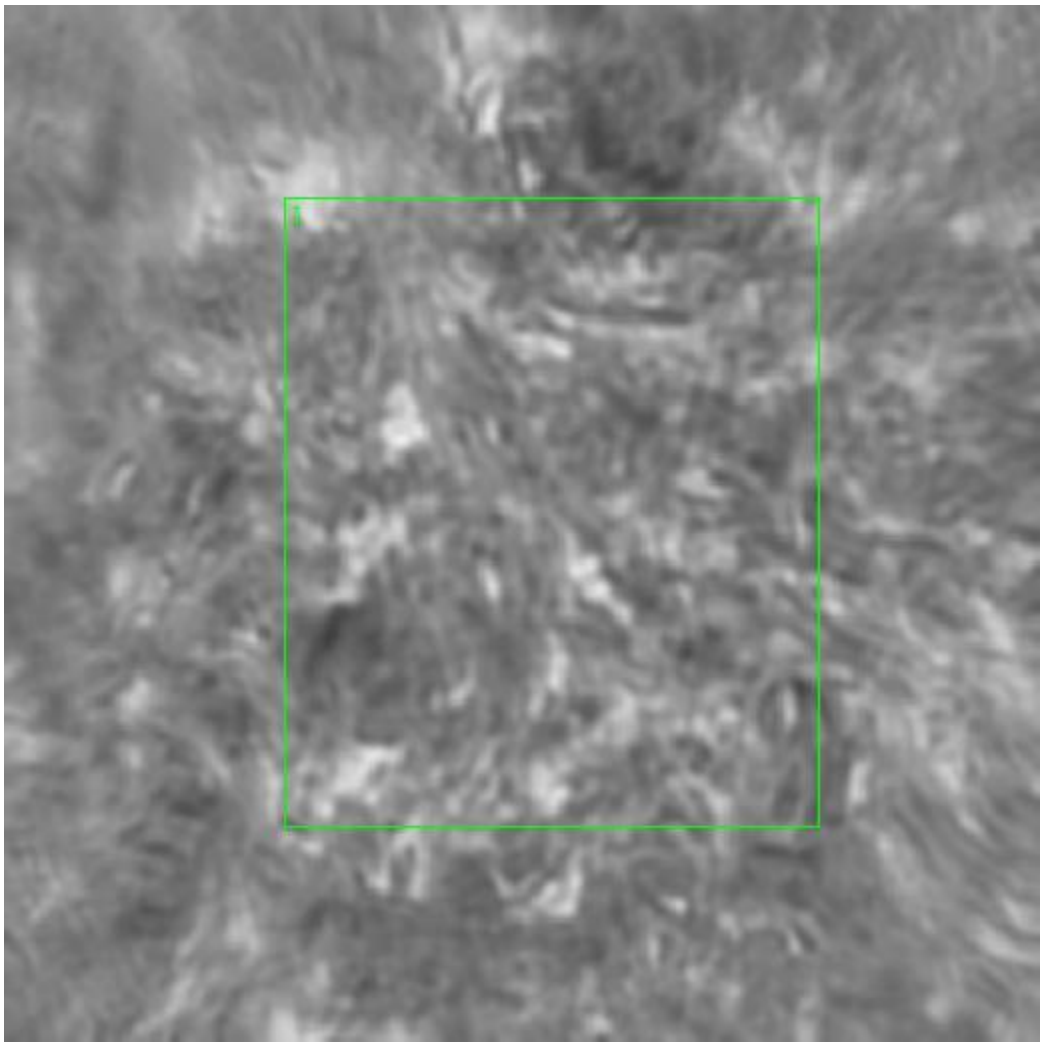


**Figure 0.46. Point P Region of Interest**
**Region 1: basement membrane, lamina propia**

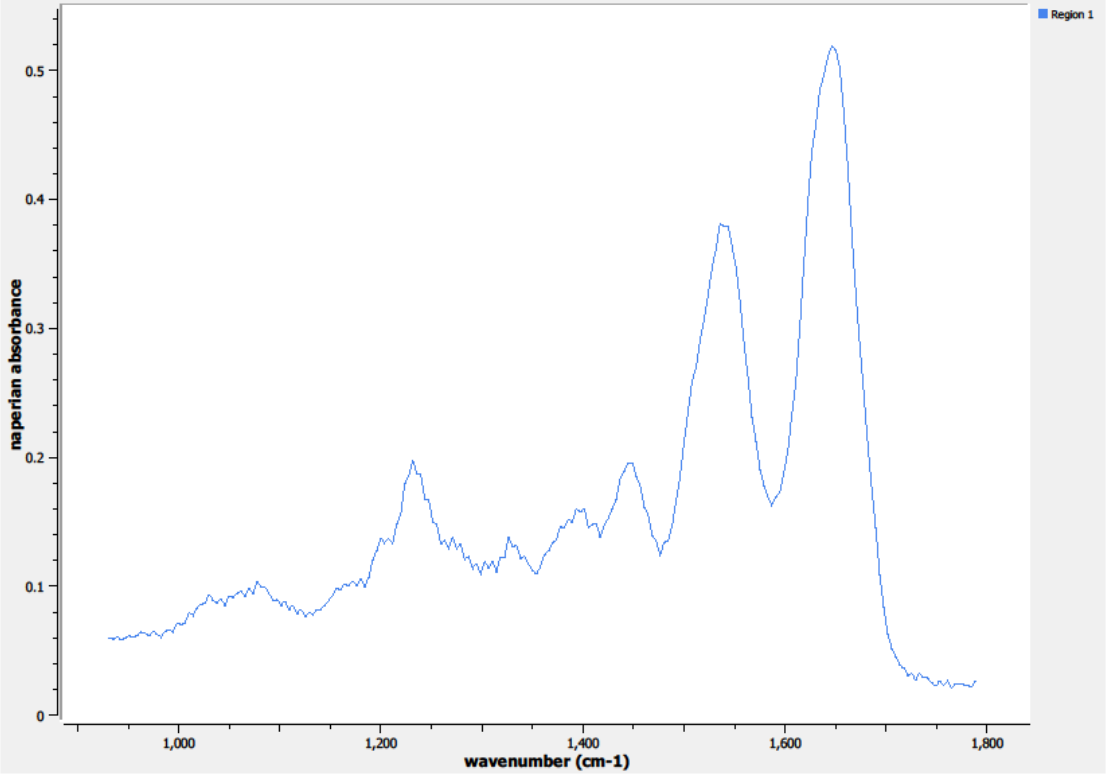



**Figure 0.47. Point P Spectra**

## A7.12-05-A043a

| | |
|---|---|
| Unique ID | 12-05-A043a |
| Age | 78 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | benign, no distinct epithelium, not normal tissue sample, unused in data results |

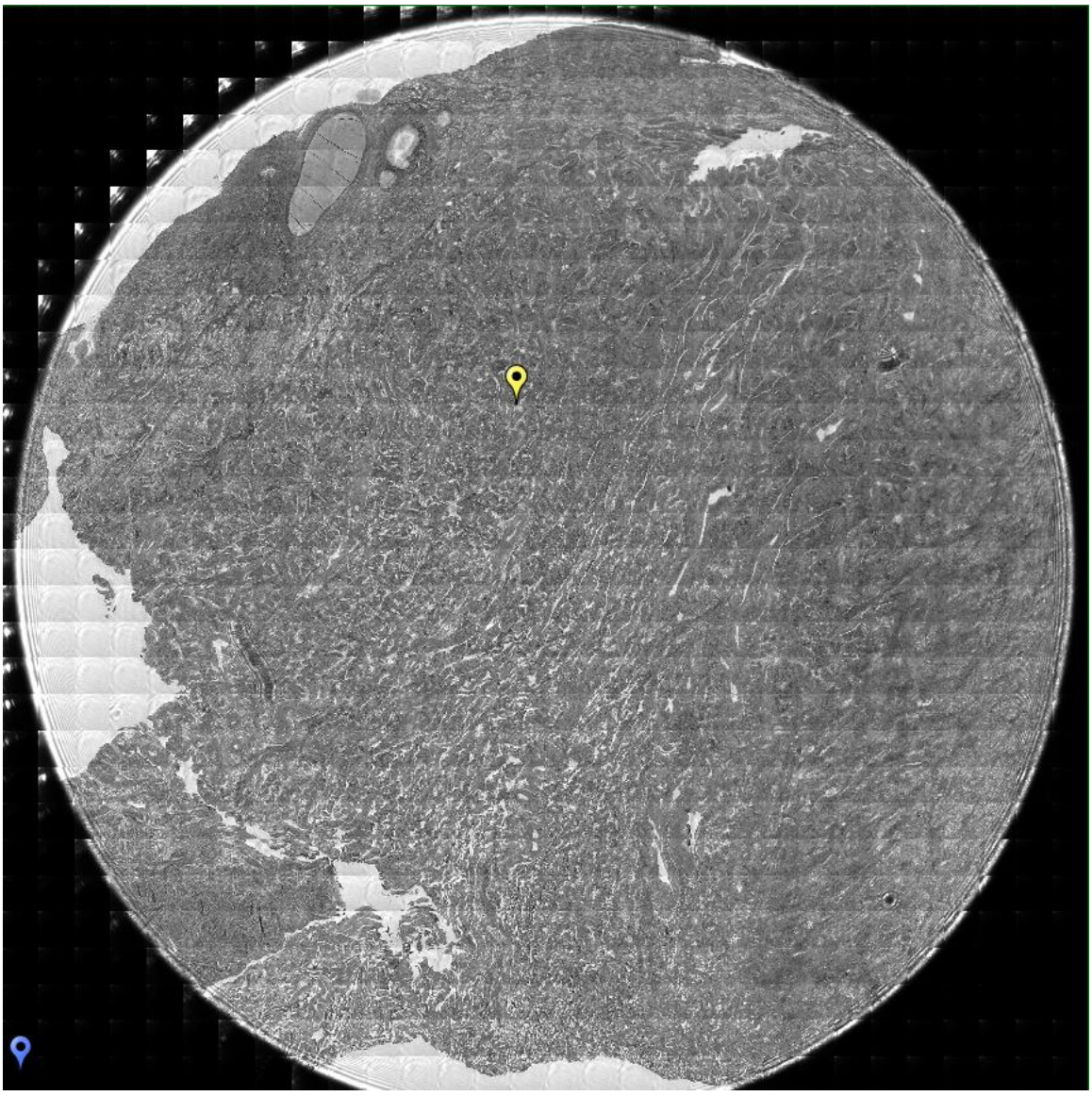


**Figure 0.48. Specimen 12-05-A043a, normal cervical tissue.**

## A8.12-05-A046a

| | |
|---|---|
| Unique ID | 12-05-A046a |
| Age | 34 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | Poor distinction between layers |

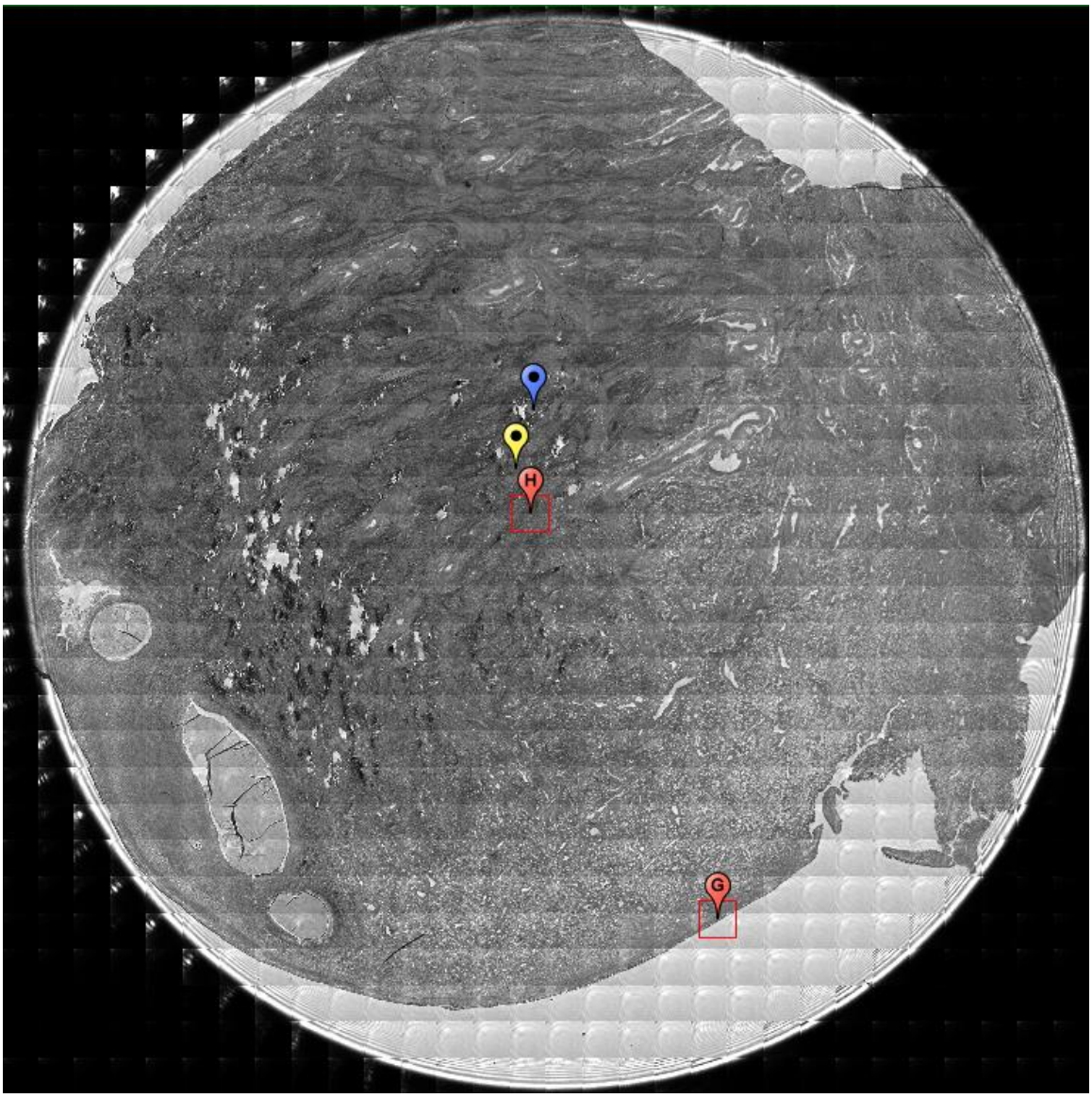



**Figure 0.49. Specimen 12-05-A046a, normal cervical tissue.**

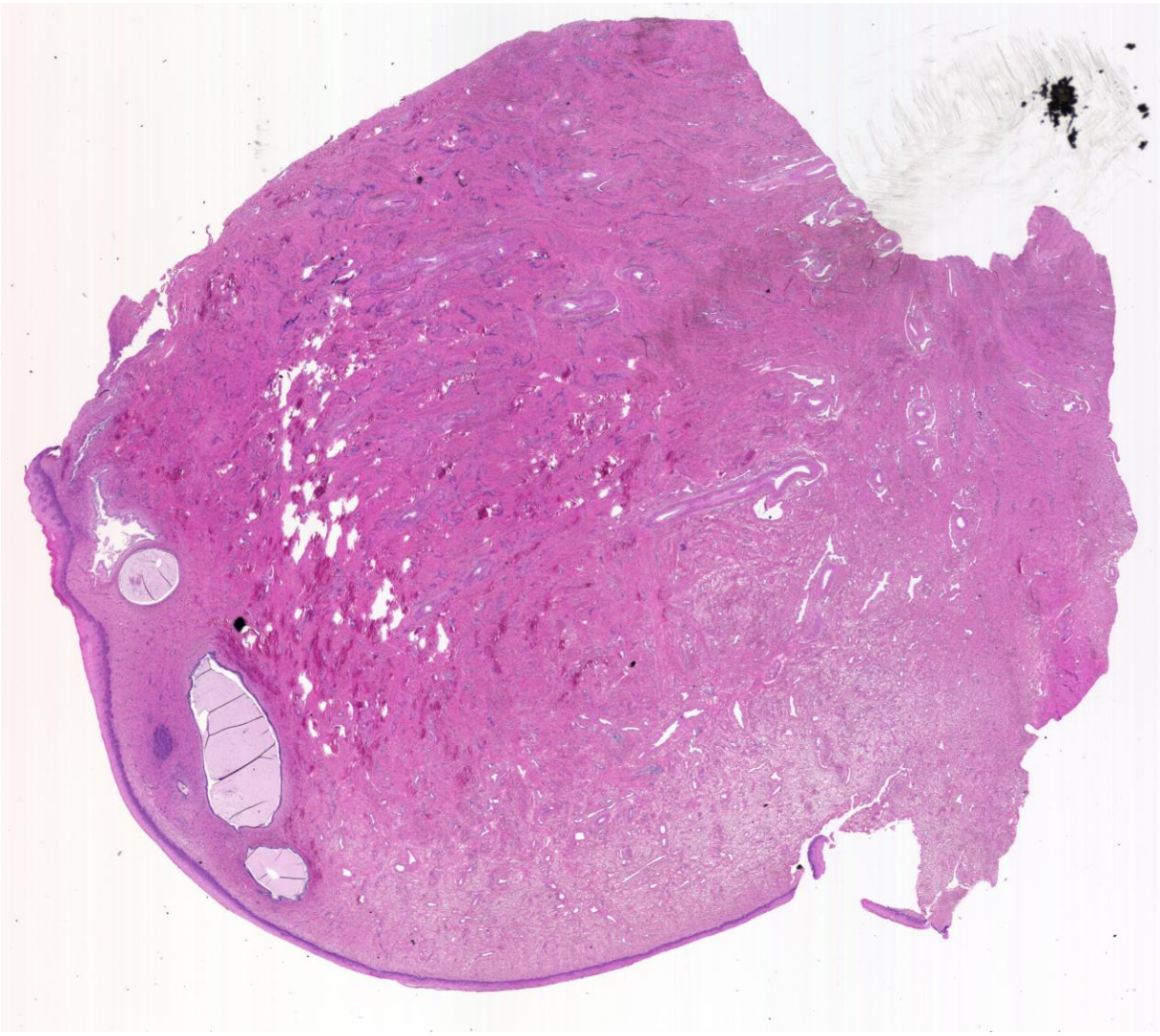


**Figure 0.50. Specimen 12-05-A046a, H&E stain, normal cervical tissue.**

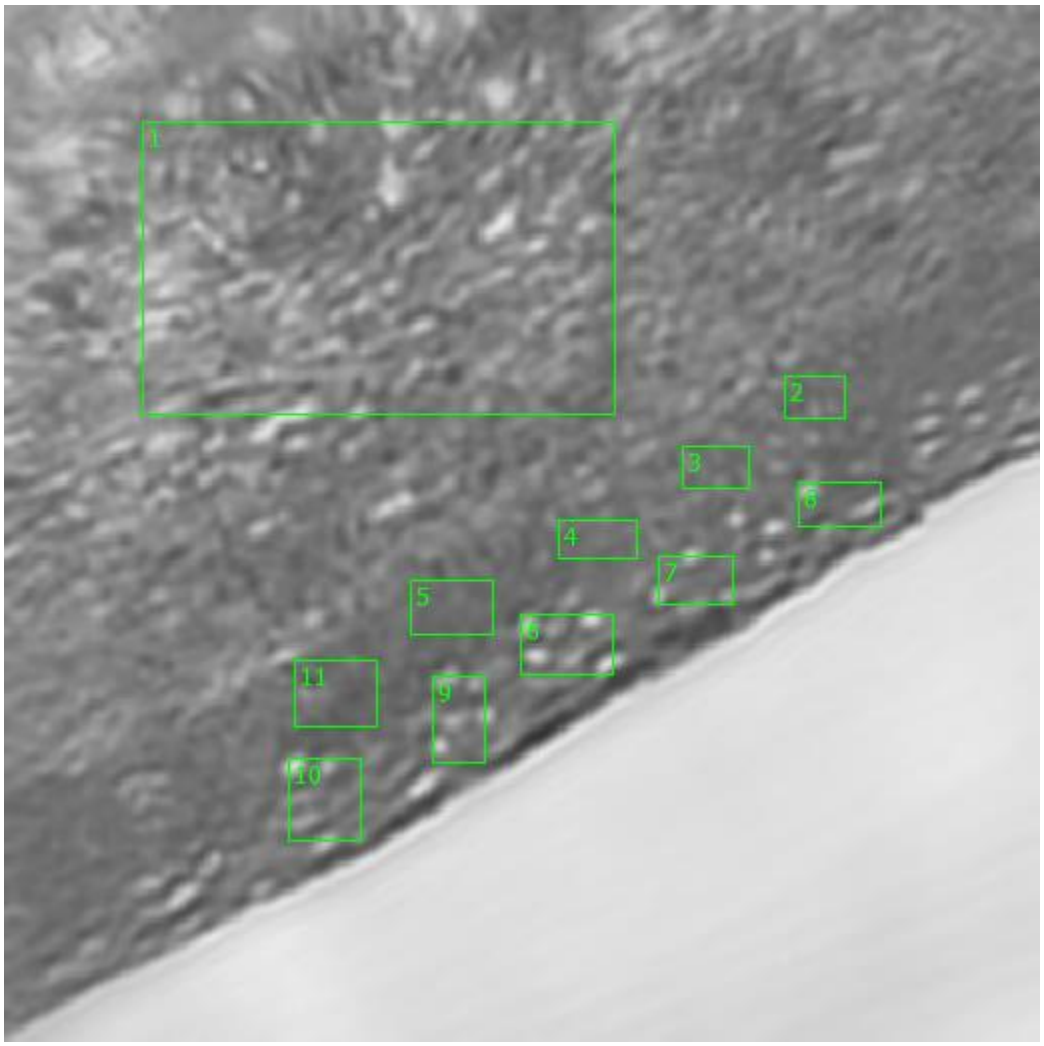



**Figure 0.51. Point G Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 3, 4, 5, 11: basal layer, parabasal layer**
**Regions 6, 7, 9: columnar epithelium, superficial layer**
**Regions 8, 10: stratified squamous epithelium, superficial layer**

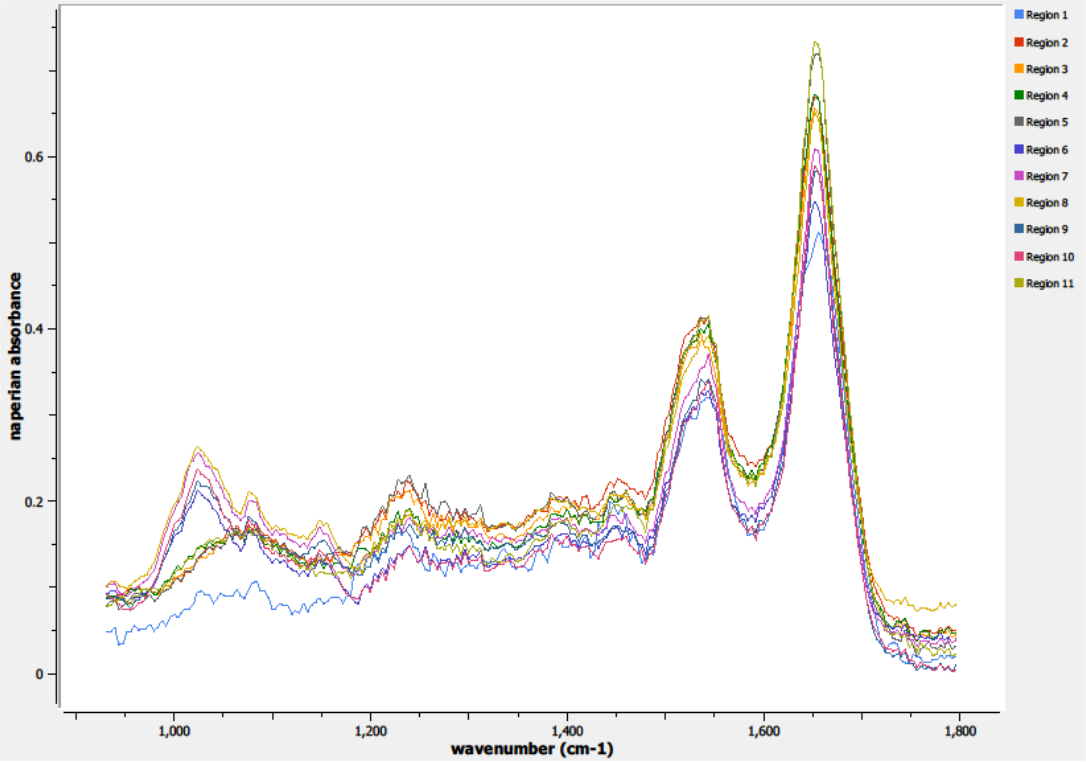



**Figure 0.52. Point G Spectra**

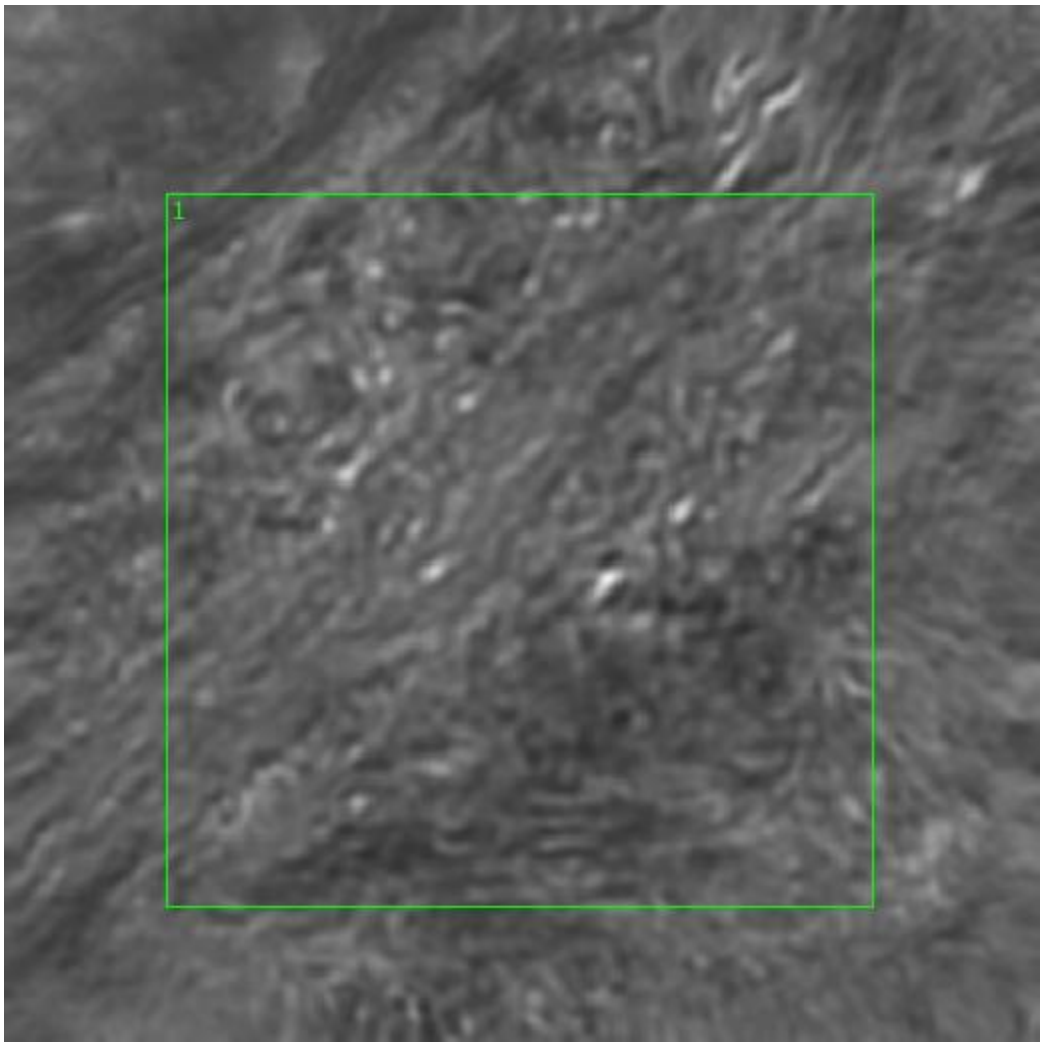



**Figure 0.53. Point H Regions of Interest**
**Region 1: lamina propia**

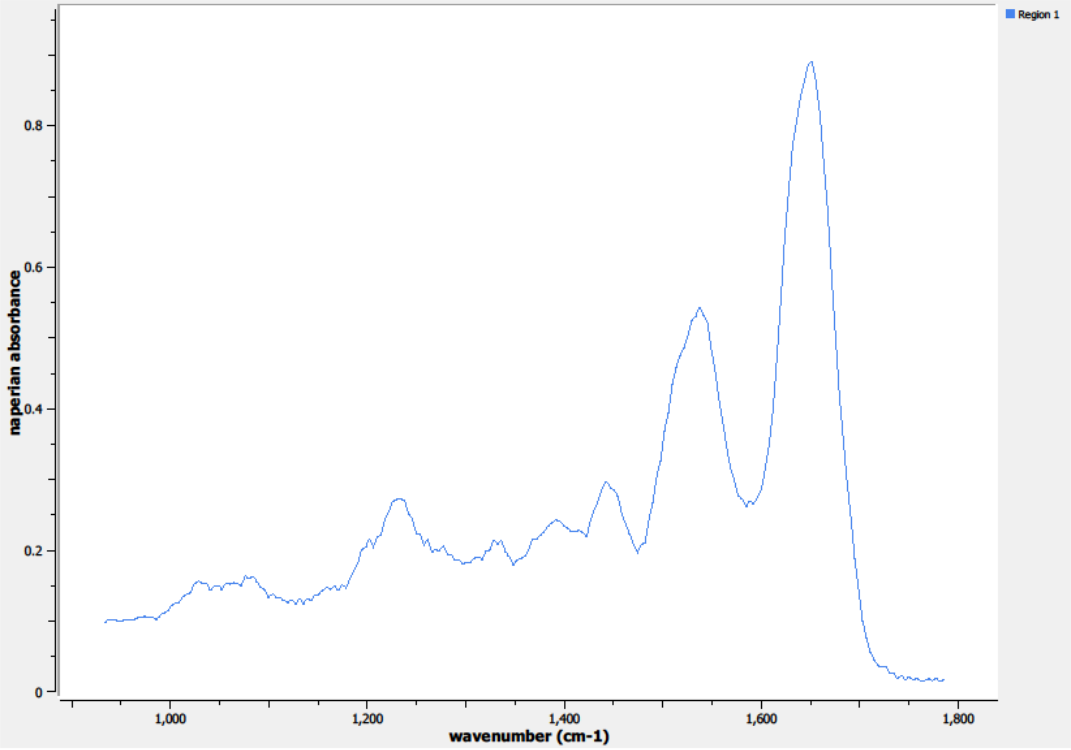



**Figure 0.54. Point H Spectra**

## A9.12-05-A072a

| | |
|---|---|
| Unique ID | 12-05-A072a |
| Age | 55 |
| Race | hispanic |
| Location of Sample Collection | ectocervix / endocervix junction |
| Notes of Interest | none |

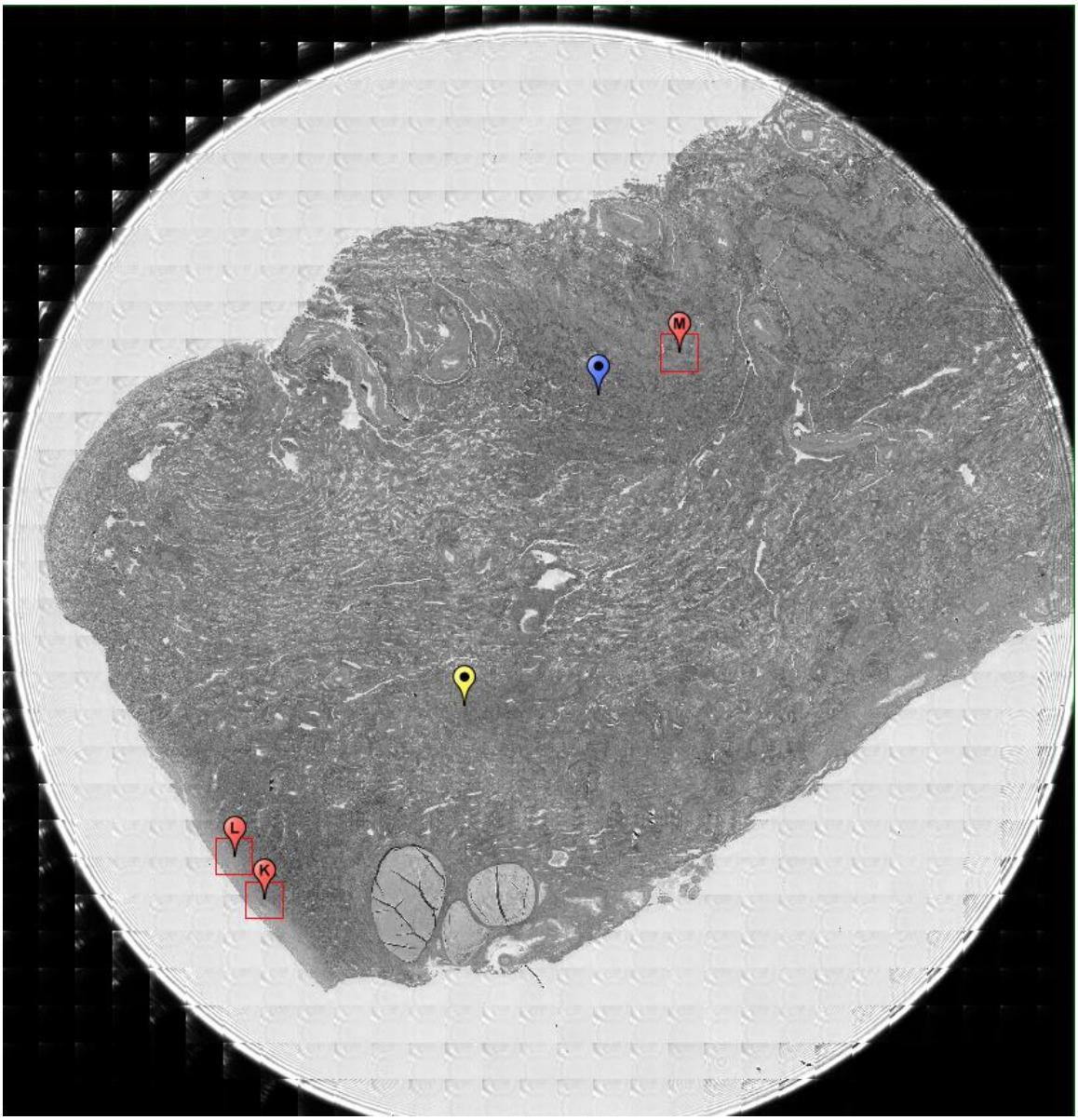



**Figure 0.55. Specimen 12-05-A072a, normal cervical tissue.**

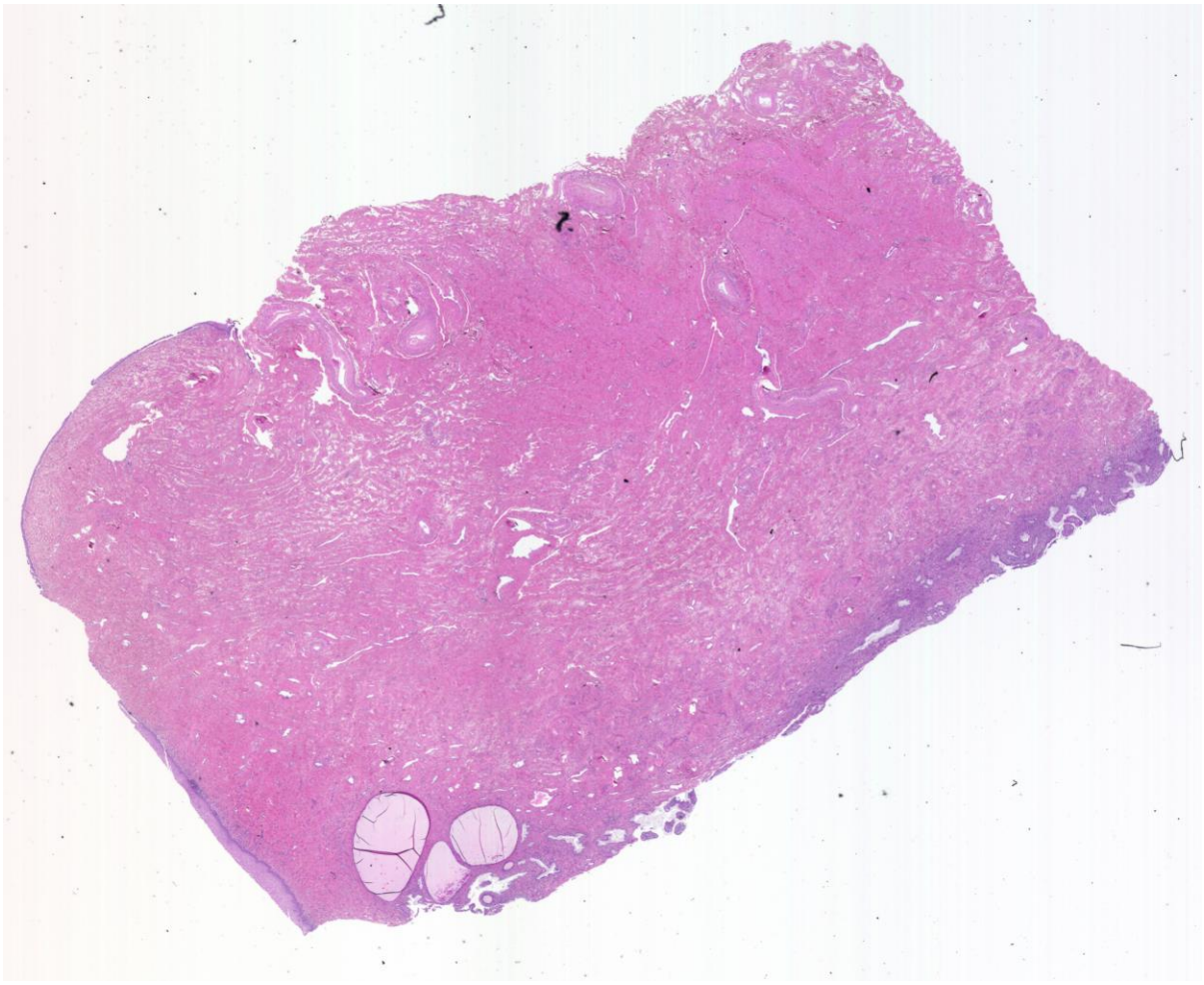


**Figure 0.56. Specimen 12-05-A072a, H&E stain, normal cervical tissue.**

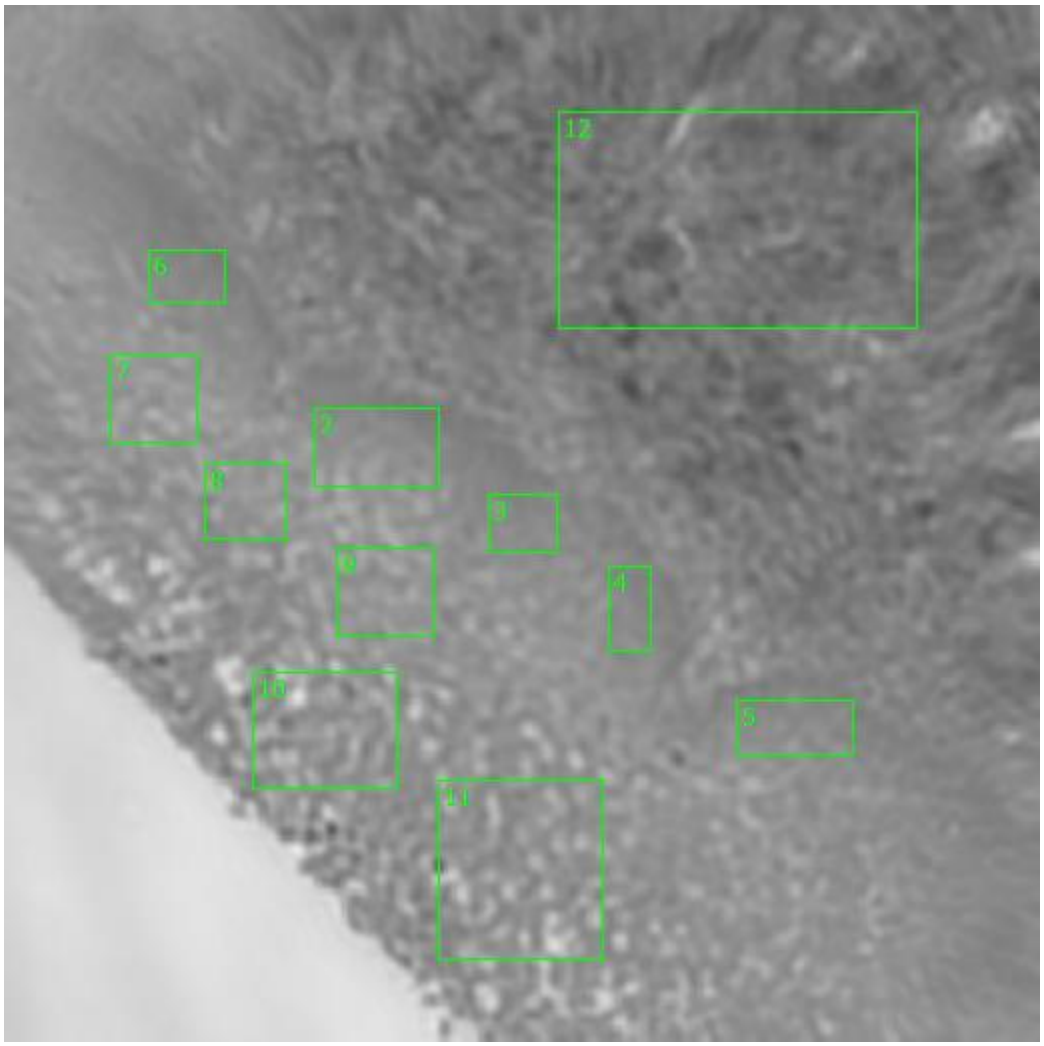



**Figure 0.57. Point K Regions of Interest**
**Regions 2, 3, 4, 5, 6: basal layer, parabasal layer**
**Regions 7, 8, 9: columnar epithelium, intermediate layer**
**Regions 10, 11: columnar epithelium, superficial layer**
**Region 12: basement membrane, lamina propia**

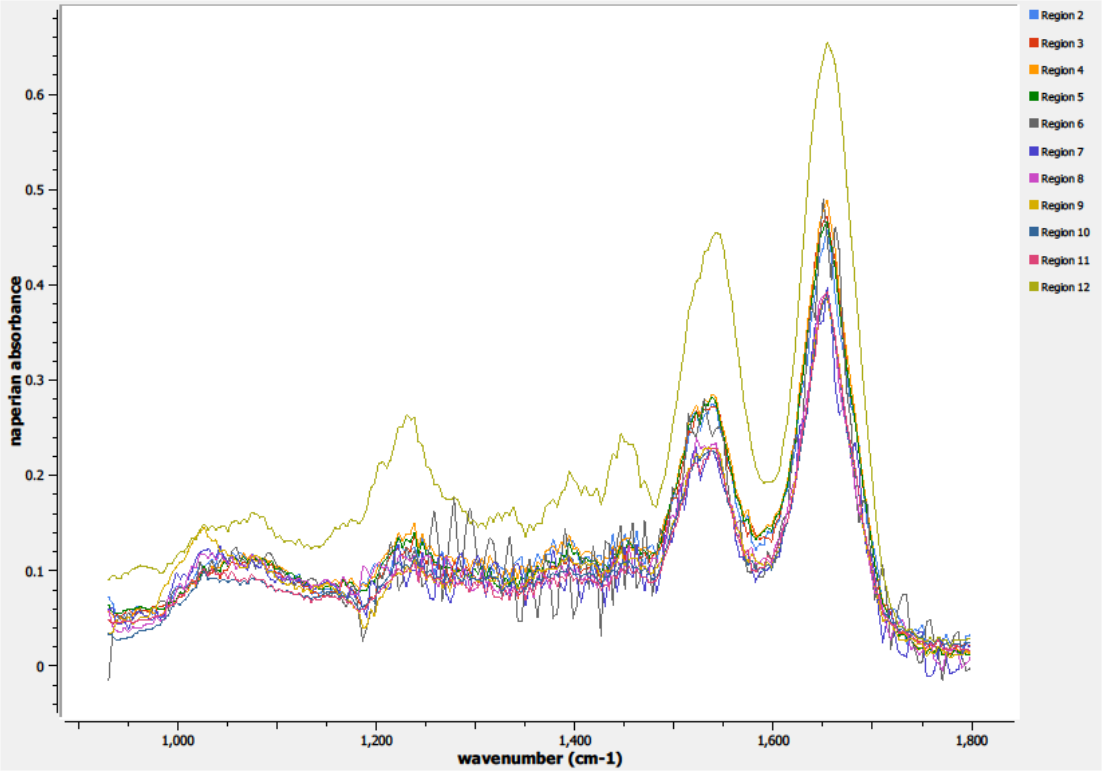



**Figure 0.58. Point K Spectra**

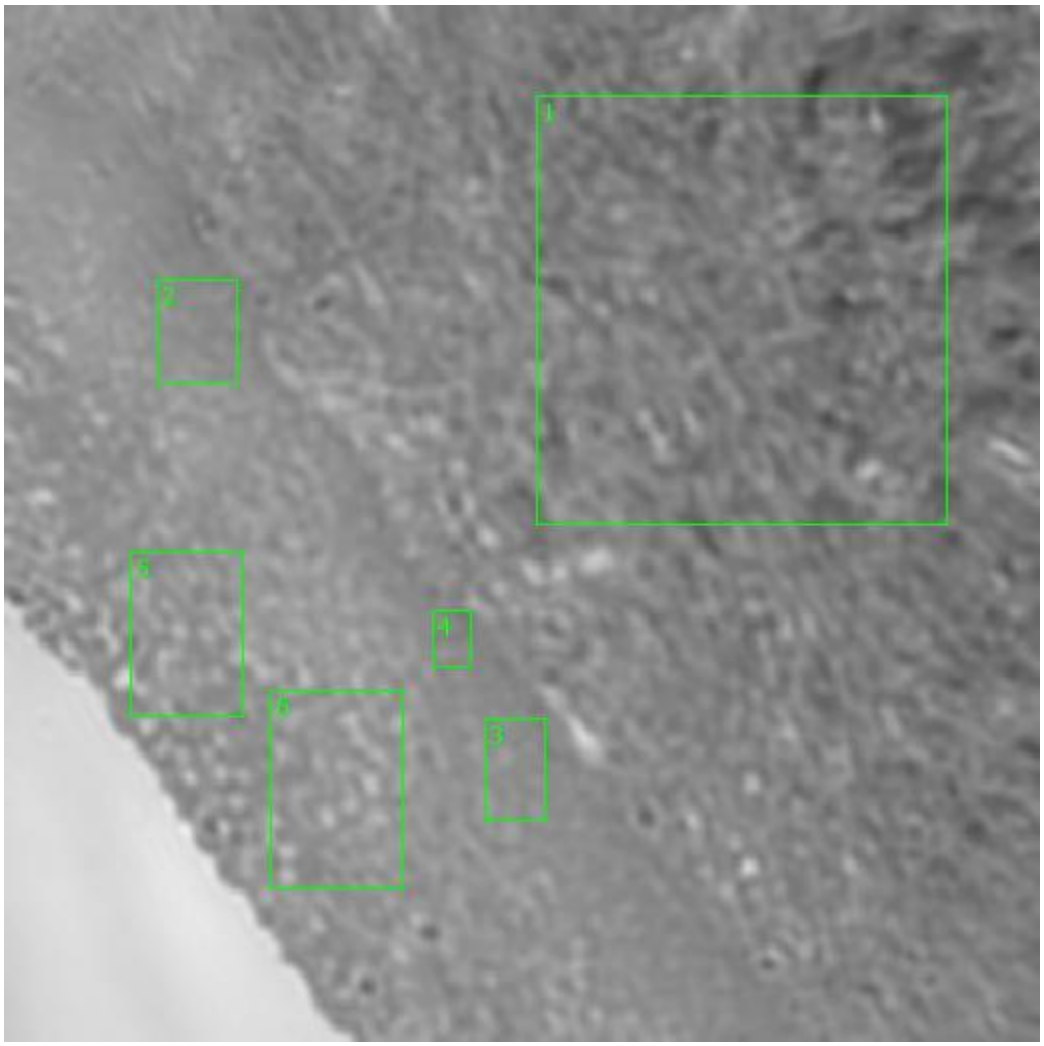



**Figure 0.59. Point L Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 3, 4: basal layer, parabasal layer**
**Regions 5, 6: columnar epithelium, superficial layer**

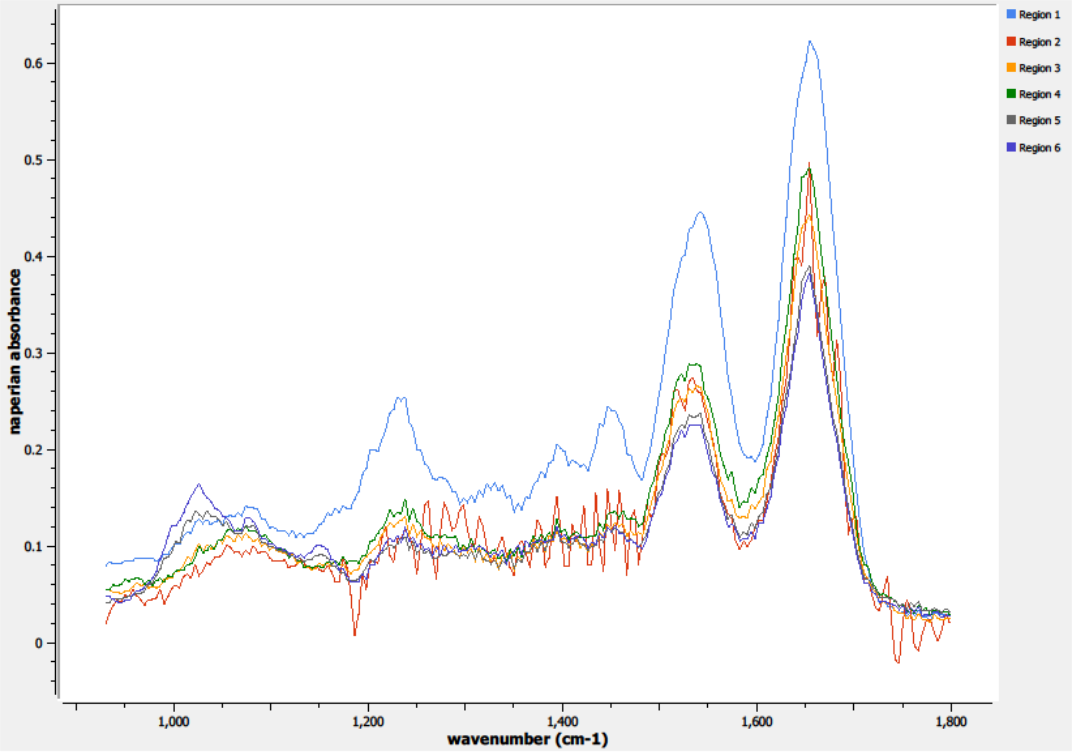



**Figure 0.60. Point L Spectra**

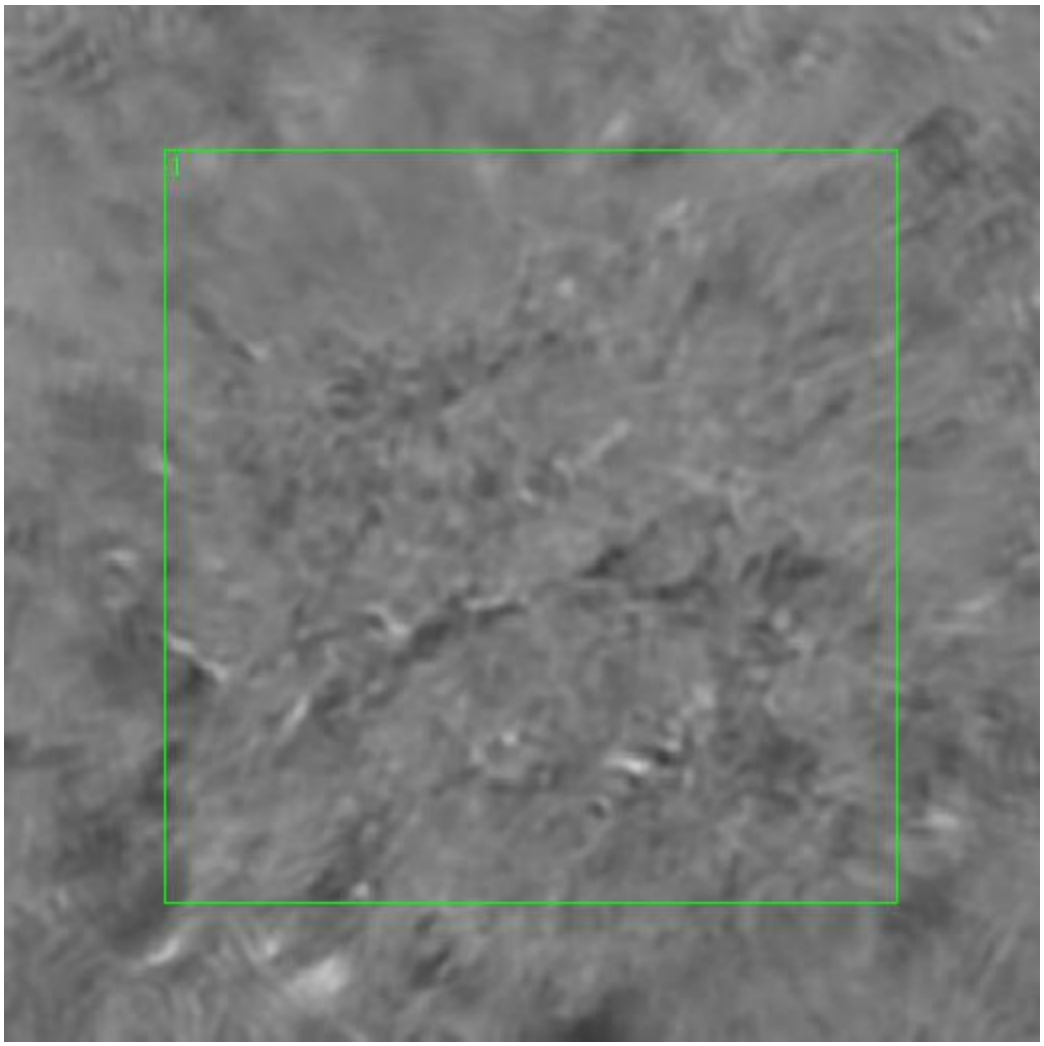


**Figure 0.61. Point M Regions of Interest**
**Region 1: lamina propia**

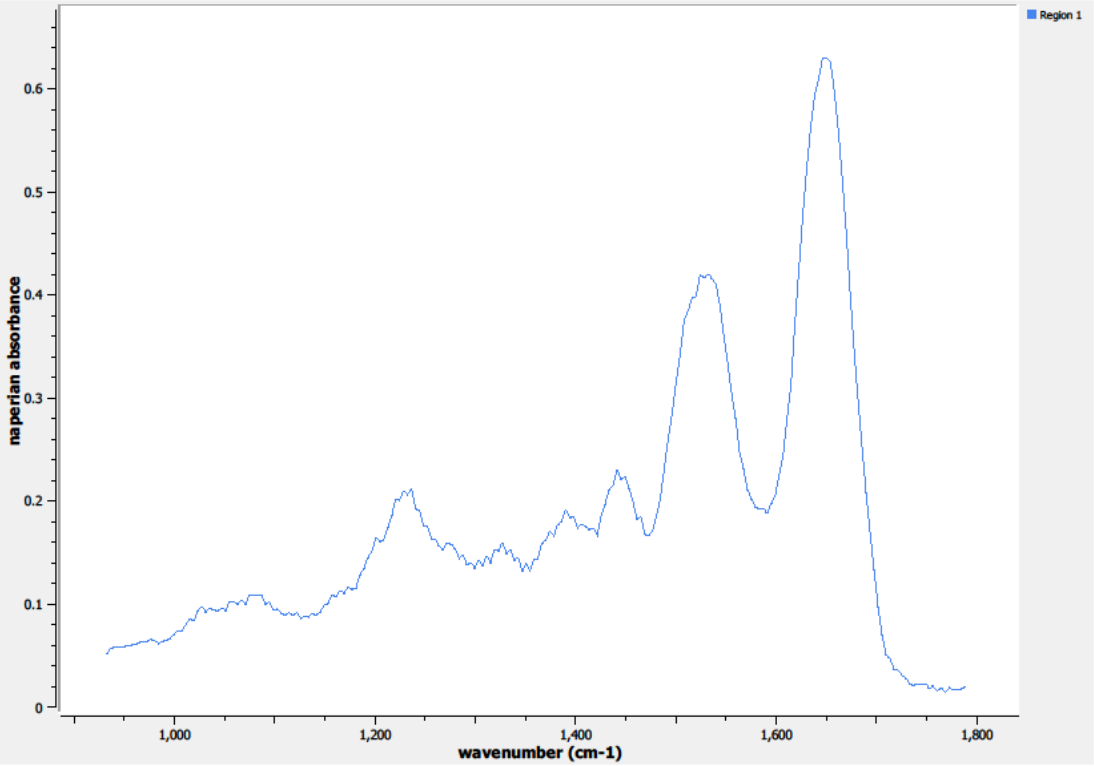



**Figure 0.62. Point M Spectra**

## A10. 12-05-A076a

| | |
|---|---|
| Unique ID | 12-05-A076a |
| Age | 59 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | none |

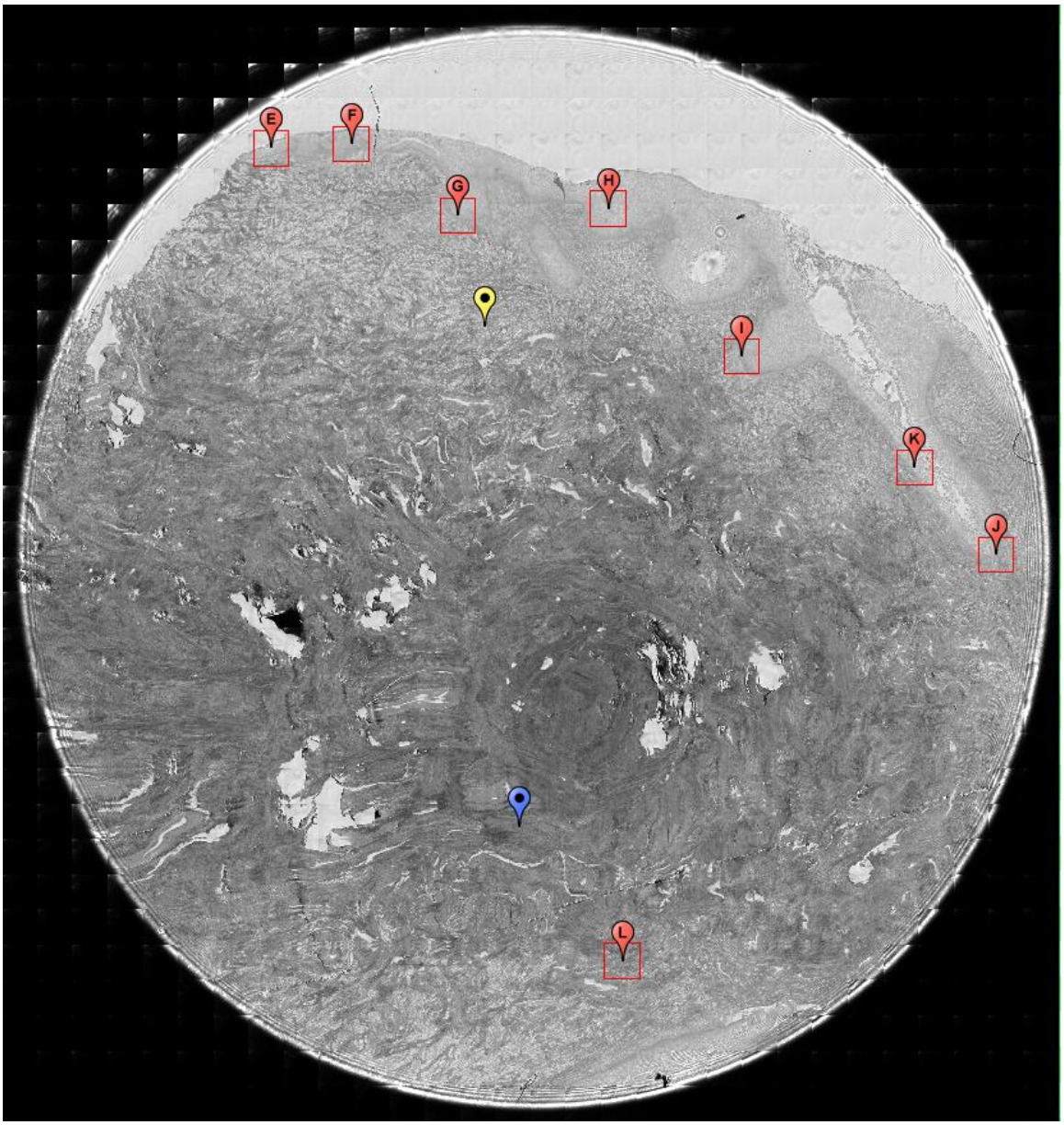



**Figure 0.63. Specimen 12-05-A076a, normal cervical tissue.**

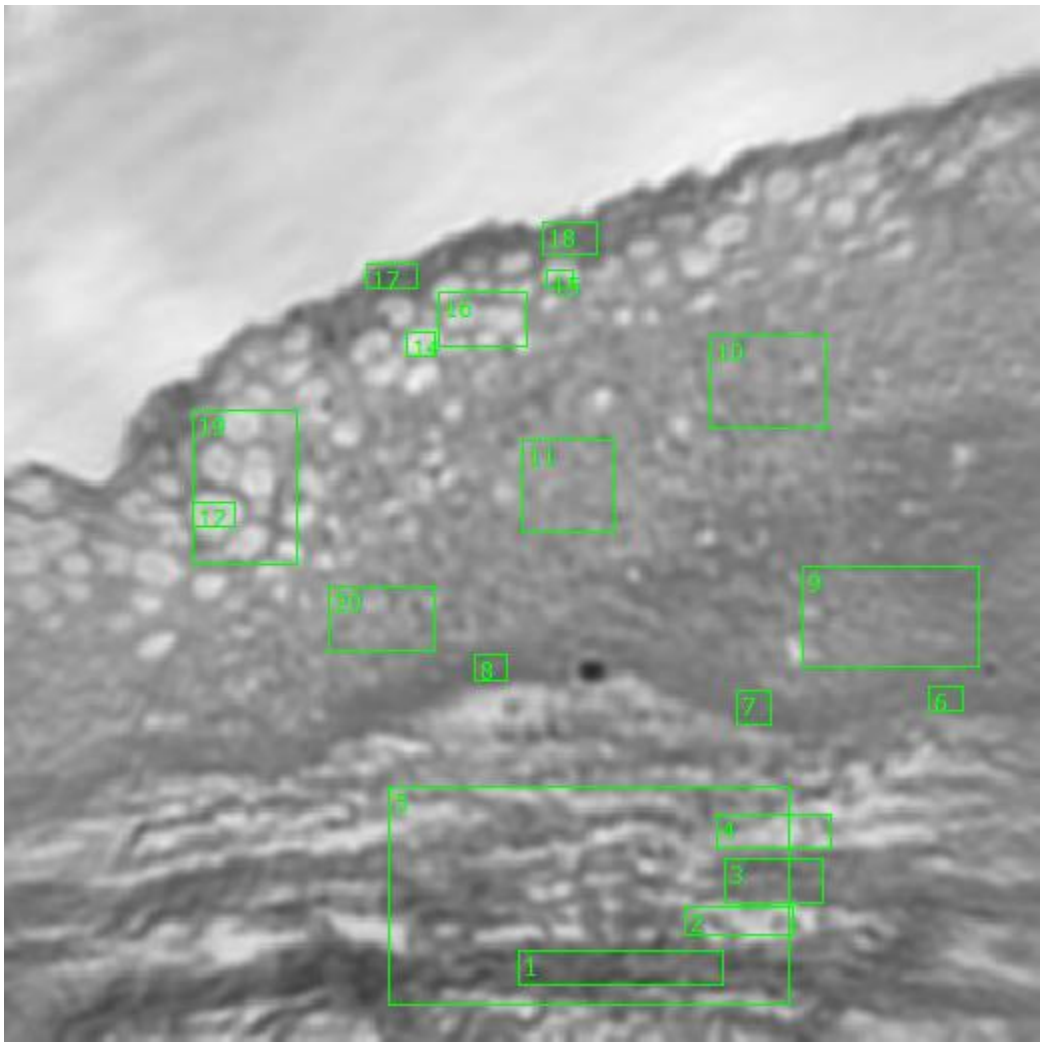



**Figure 0.64. Point E Regions of Interest**
**Regions 1, 2, 3, 4, 5: basement membrane, lamina propia**
**Regions 6, 7, 8: basal layer**
**Region 9: parabasal layer**
**Regions 10, 11, 20: columnar epithelium**
**Regions 12, 14, 15: individual columnar epithelial cells**
**Regions 16, 19: columnar epithelium, intermediate layer**
**Regions 17, 18: stratified squamous epithelium, epithelial surface**

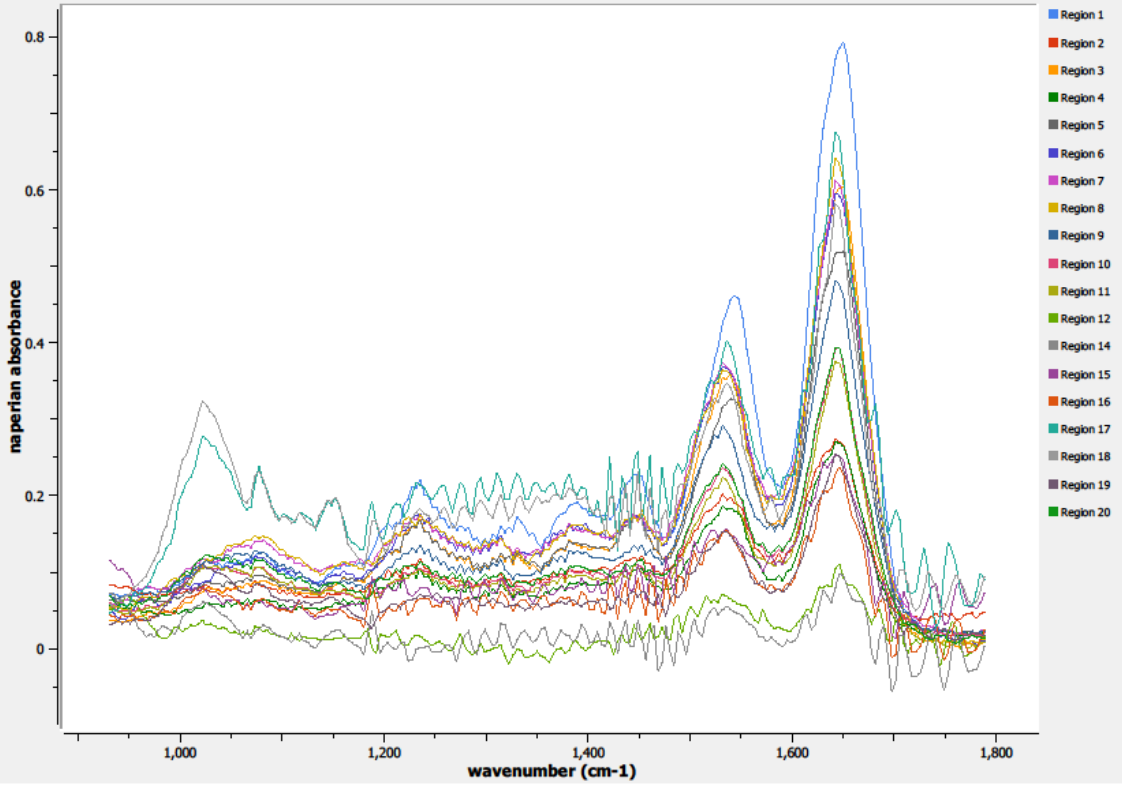



**Figure 0.65. Point E Spectra**

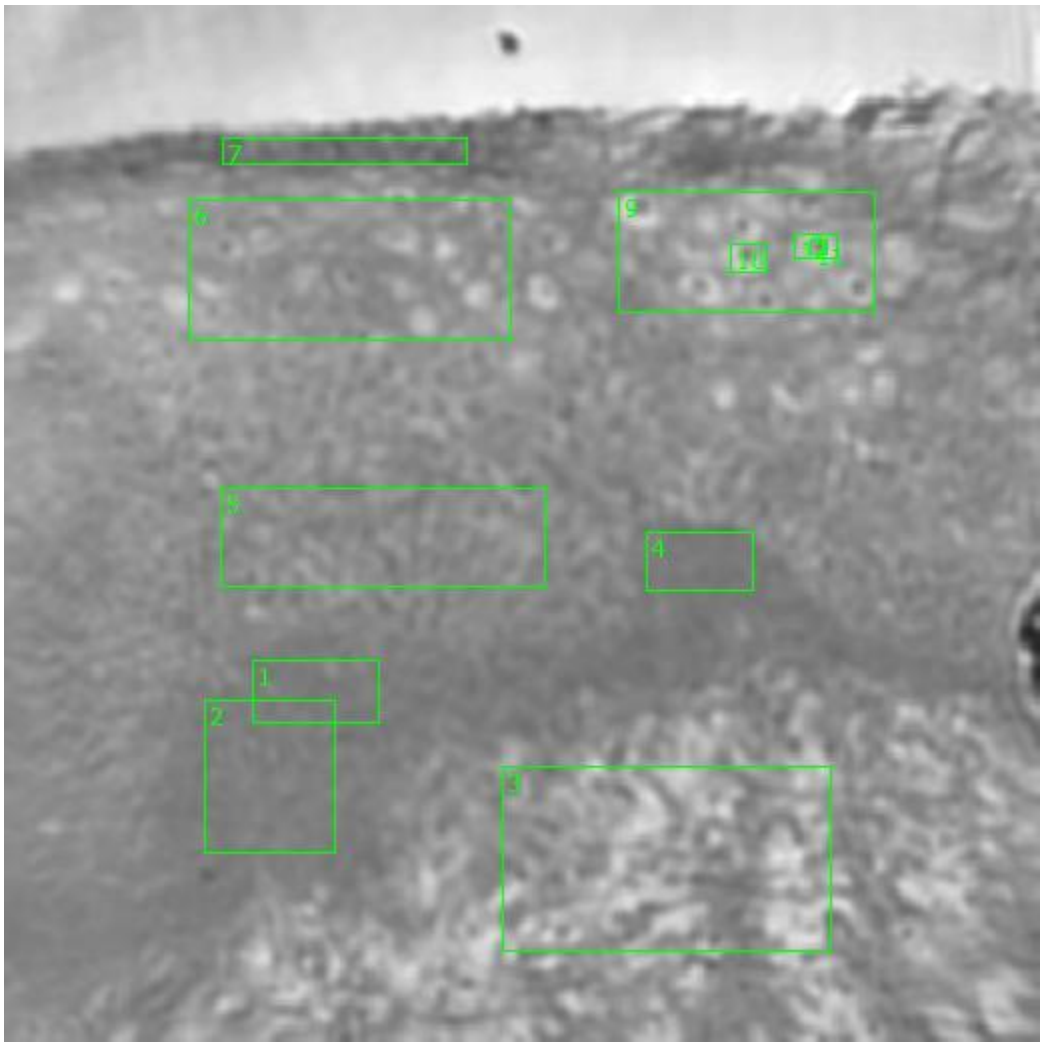



**Figure 0.66. Point F Regions of Interest**
**Regions 1, 2, 4: basal layer**
**Region 3: basement membrane, lamina propia**
**Region 5: parabasal layer**
**Regions 6, 9: columnar epithelium**
**Region 7: stratified squamous epithelium, epithelial surface**
**Regions 10, 11, 12: individual columnar epithelial cells**

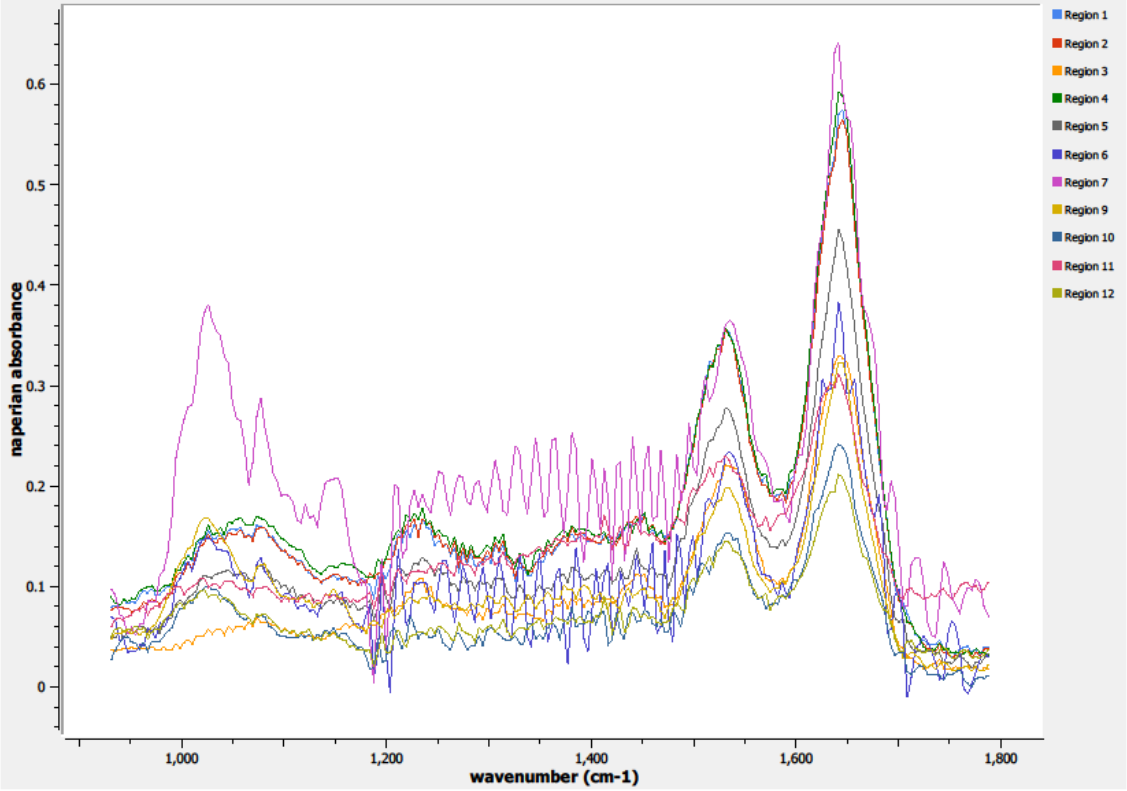



**Figure 0.67. Point F Spectra**

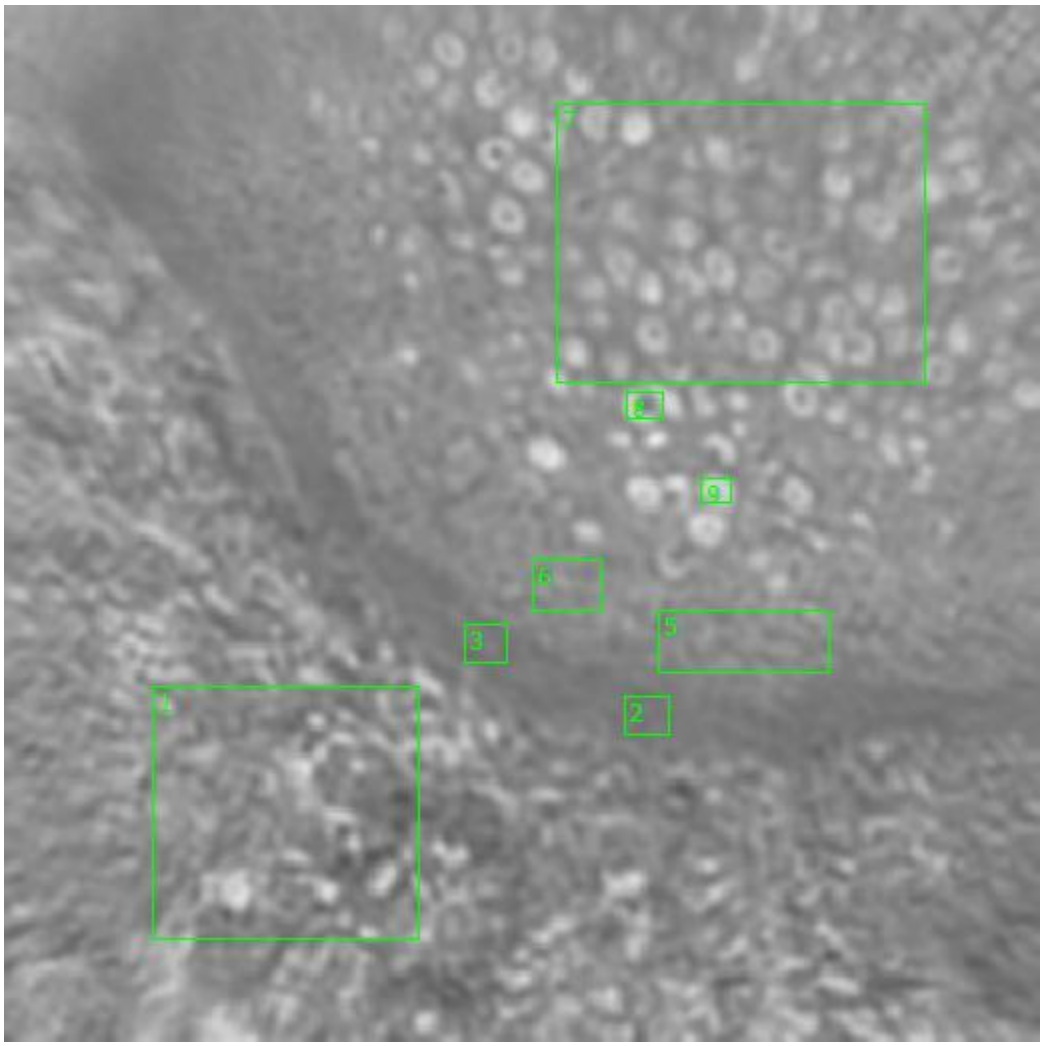



**Figure 0.68. Point G Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 3: basal layer**
**Regions 5, 6: parabasal layer**
**Region 7: columnar epithelium**
**Regions 8, 9: individual columnar epithelial cells**

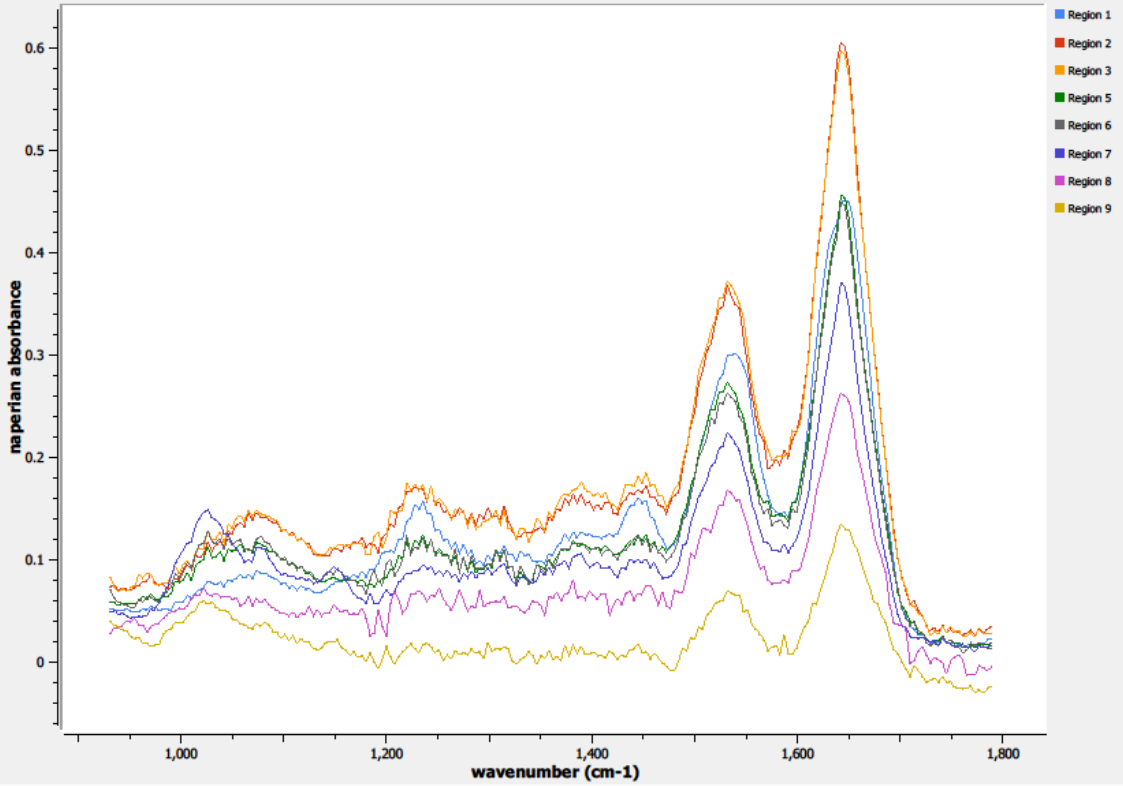



**Figure 0.69. Point G Spectra**

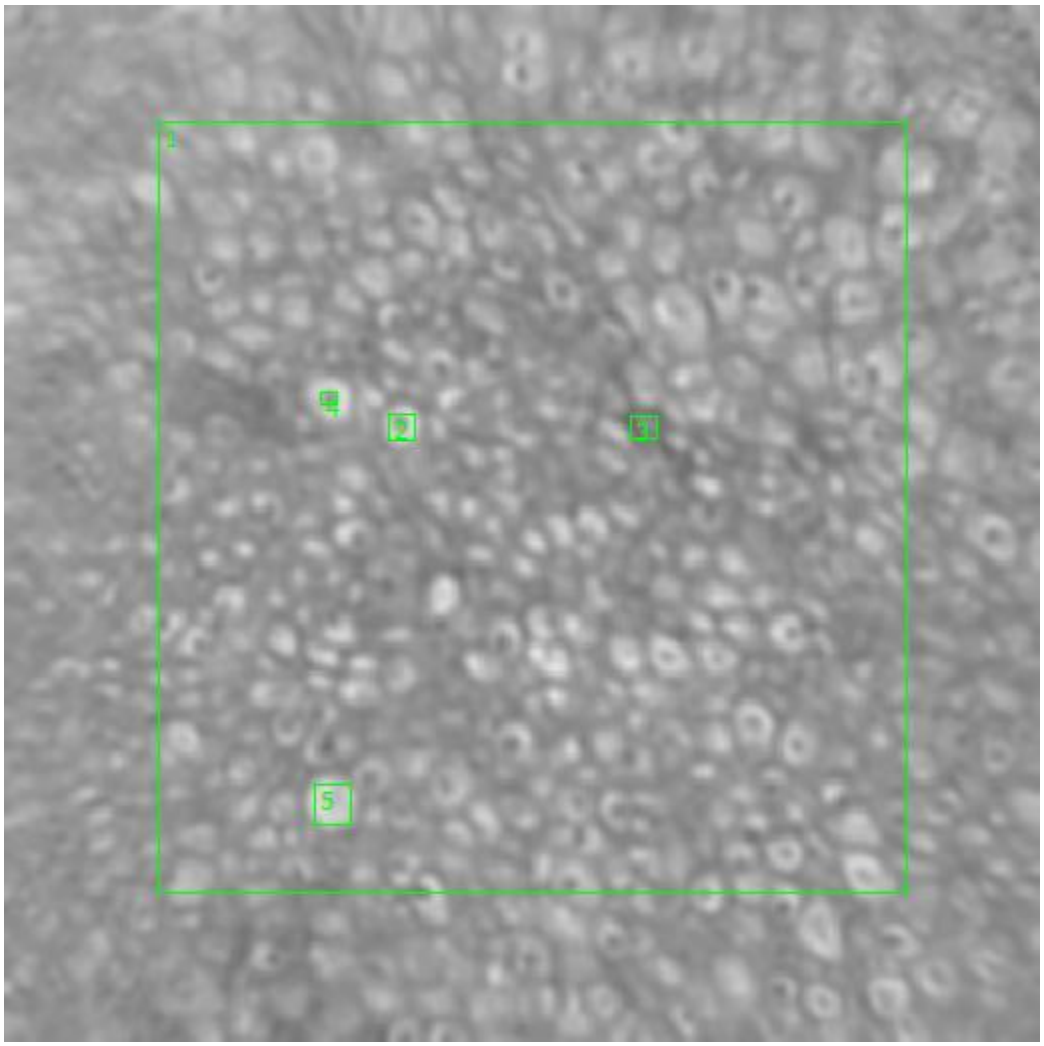



**Figure 0.70. Point H Regions of Interest**
**Regions 1, 3: columnar epithelium**
**Regions 2, 5: individual columnar epithelial cells**
**Region 4: individual columnar epithelial cell nuclei**

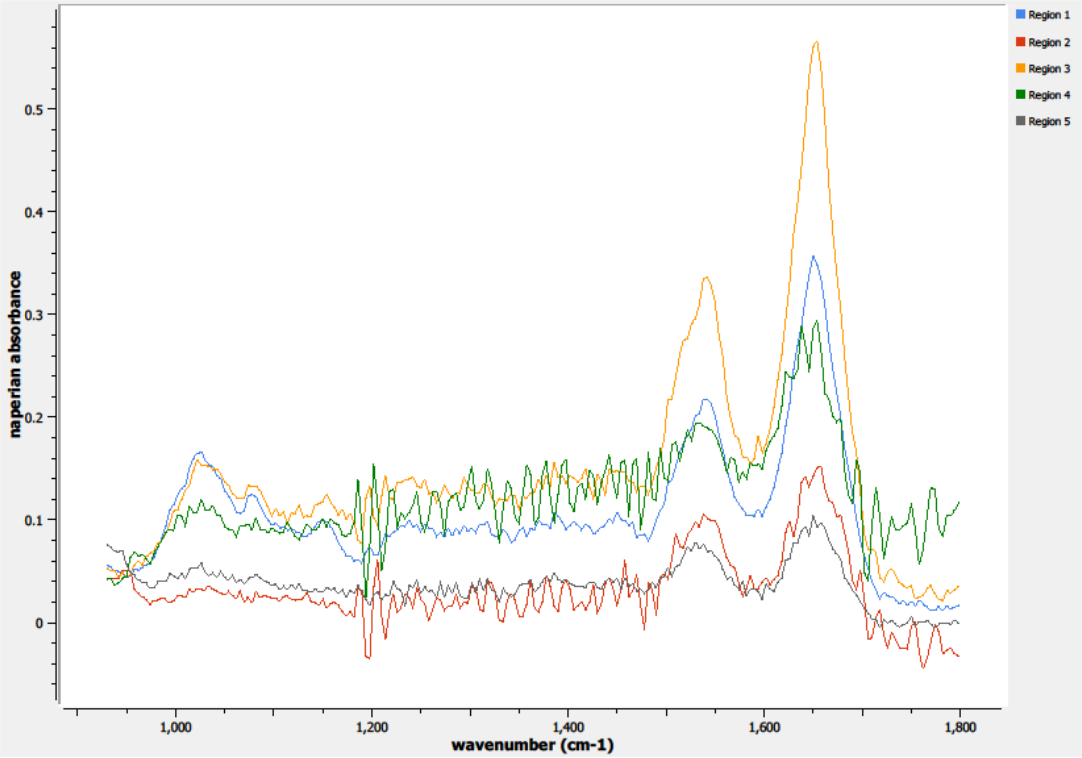



**Figure 0.71. Point H Spectra**

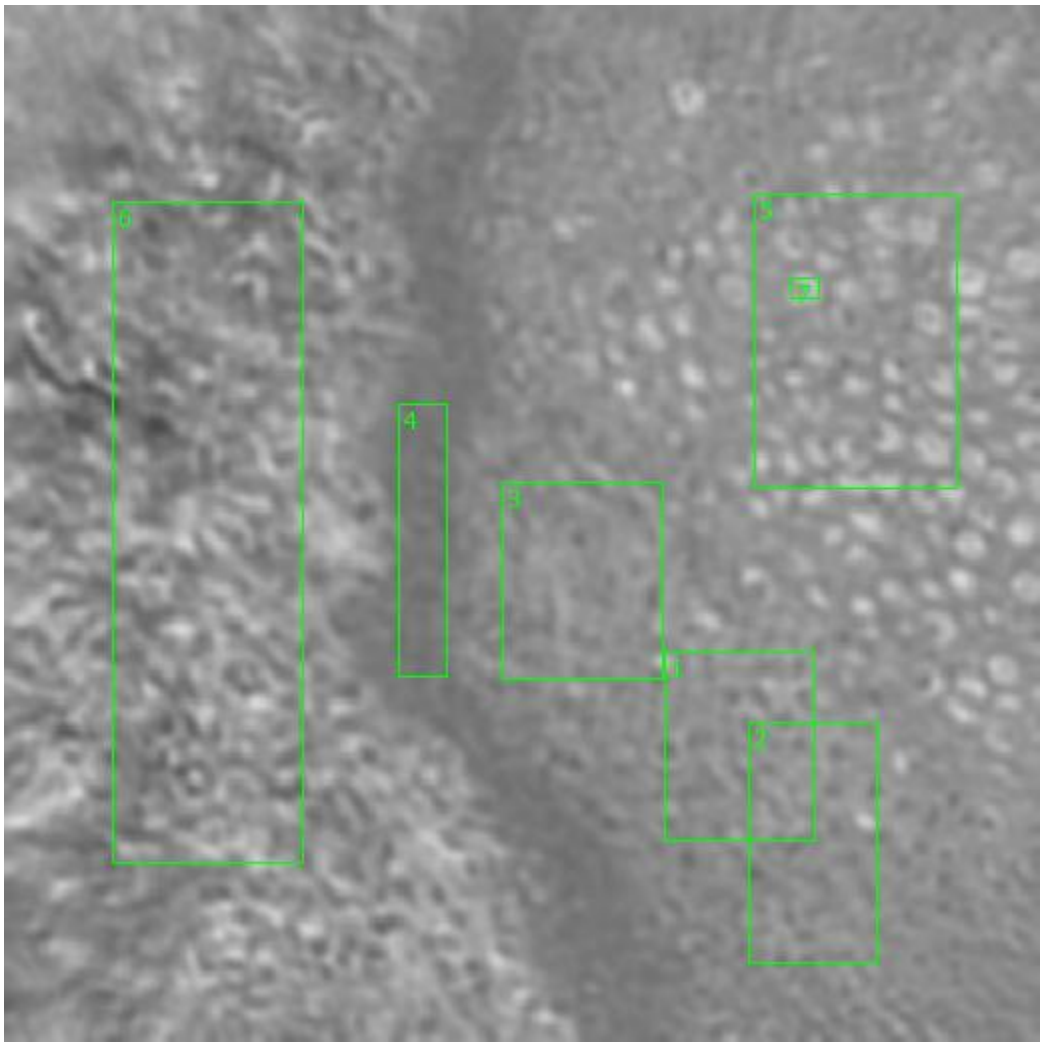



**Figure 0.72. Point I Regions of Interest**
**Regions 1, 2, 3: parabasal layer**
**Region 4: basal layer**
**Region 5: columnar epithelium**
**Region 6: basement membrane, lamina propia**
**Region 7: individual columnar epithelial cell**

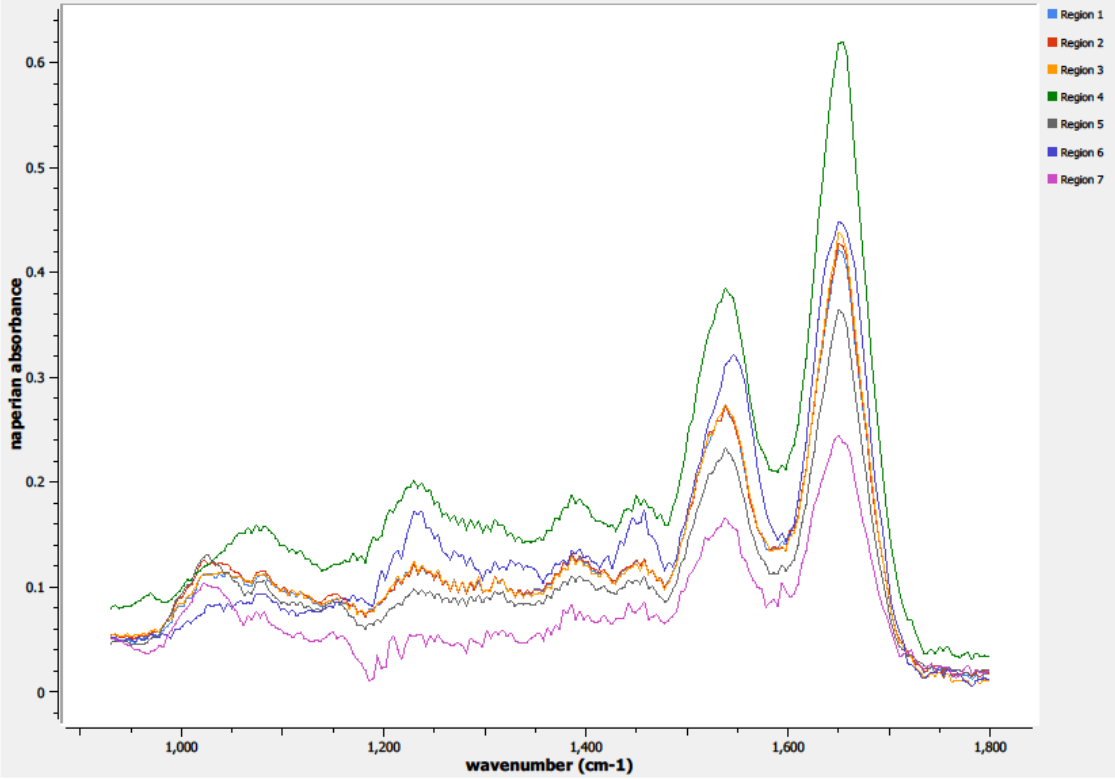



**Figure 0.73. Point I Spectra**

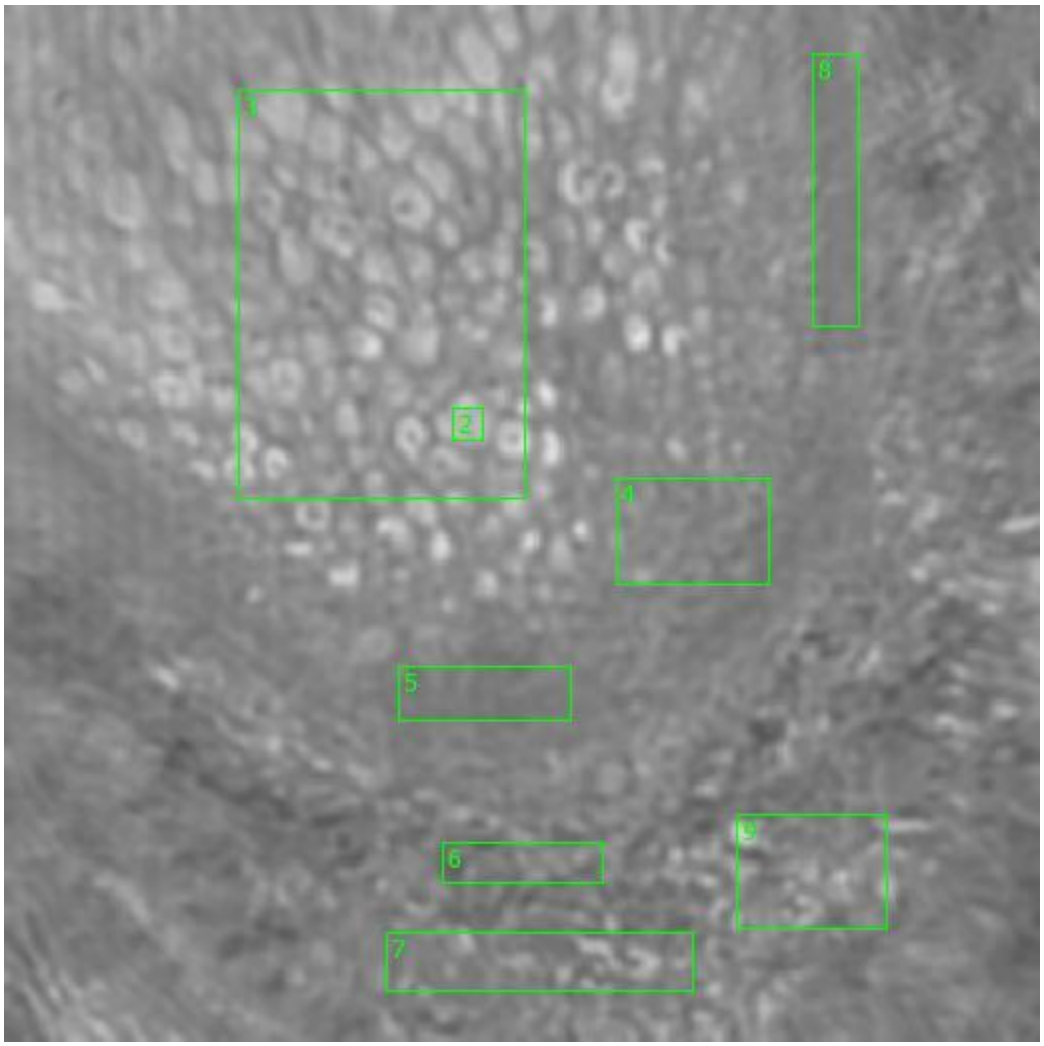



**Figure 0.74. Point J Regions of Interest**
**Region 2: individual columnar epithelial cell**
**Region 3: columnar epithelium**
**Regions 4, 5, 8: parabasal layer**
**Regions 6, 7, 9: basement membrane, lamina propia**

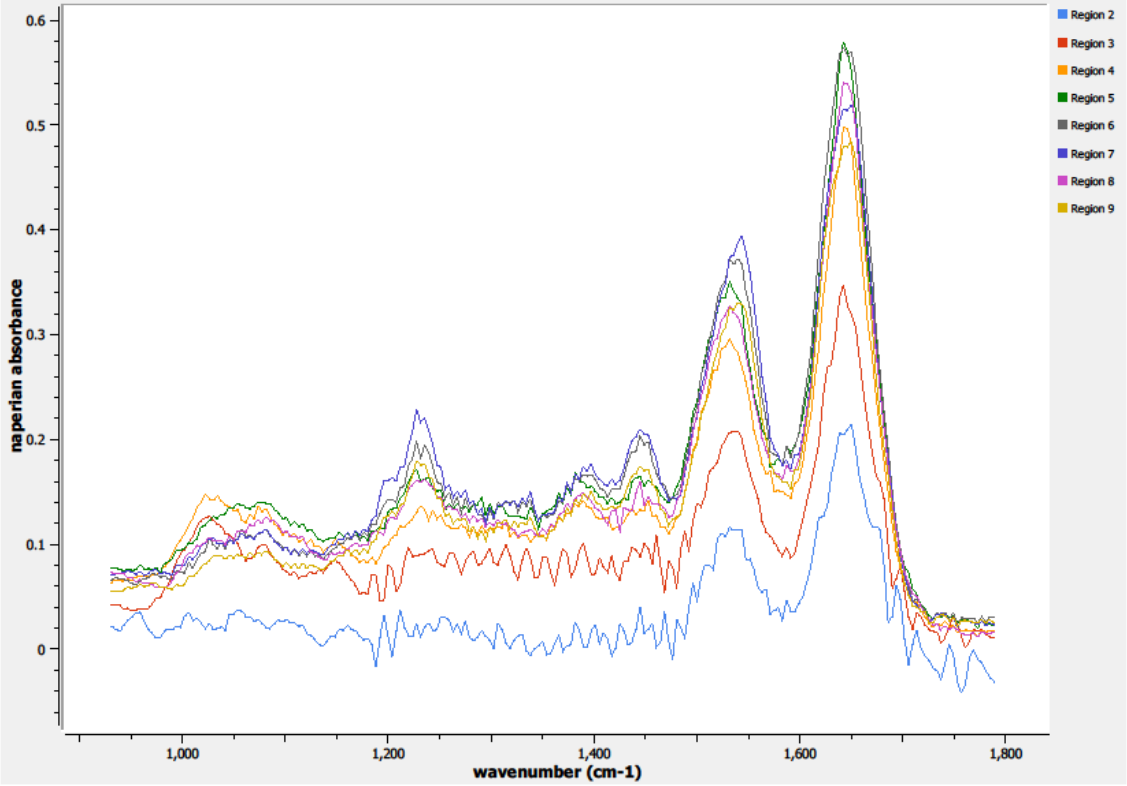



**Figure 0.75. Point J Spectra**

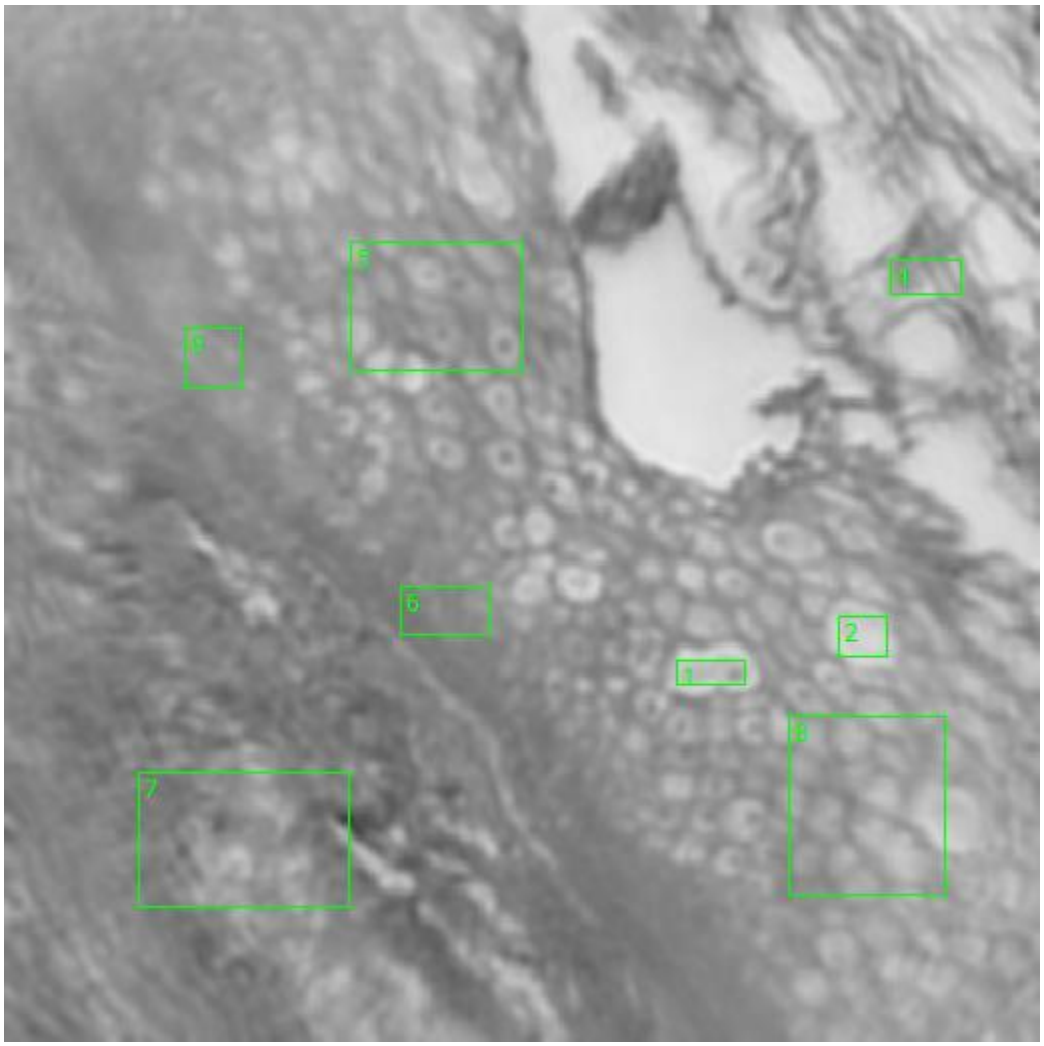



**Figure 0.76. Point K Regions of Interest**
**Regions 1, 2: individual columnar epithelial cell**
**Region 4: stratified squamous epithelium**
**Regions 5, 8: columnar epithelium**
**Regions 6, 9: basal, parabasal layer**
**Region 7: basement membrane, lamina propia**

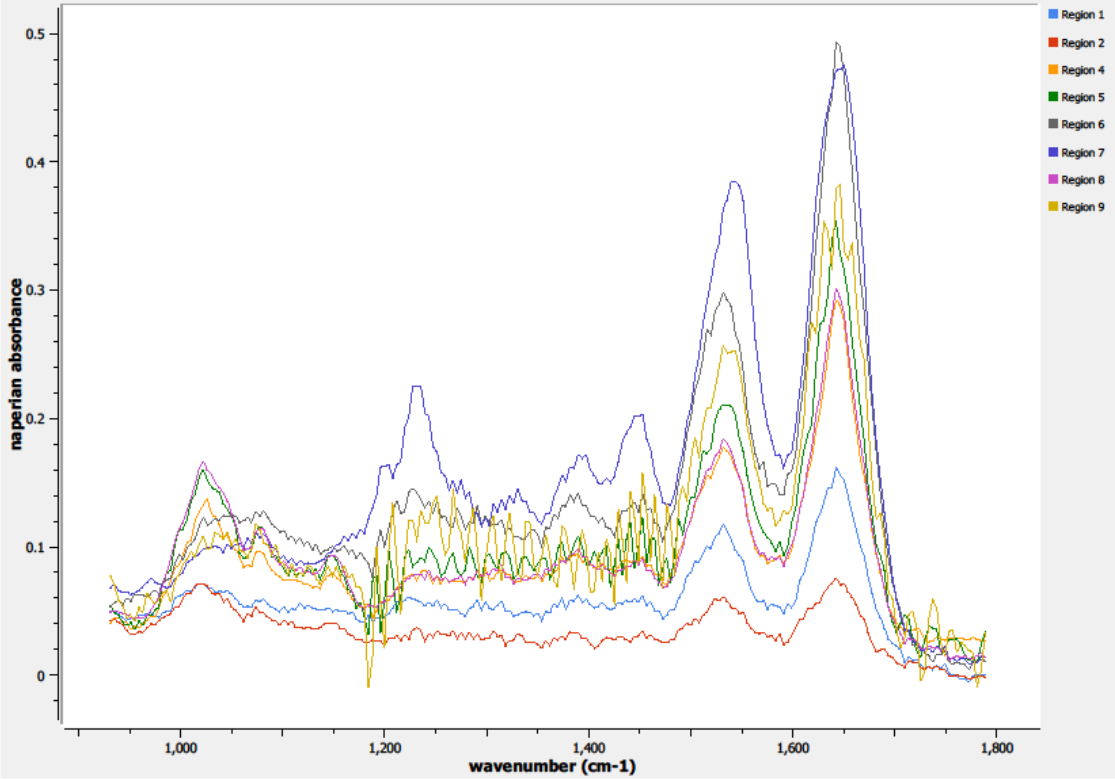



**Figure 0.77. Point K Spectra**

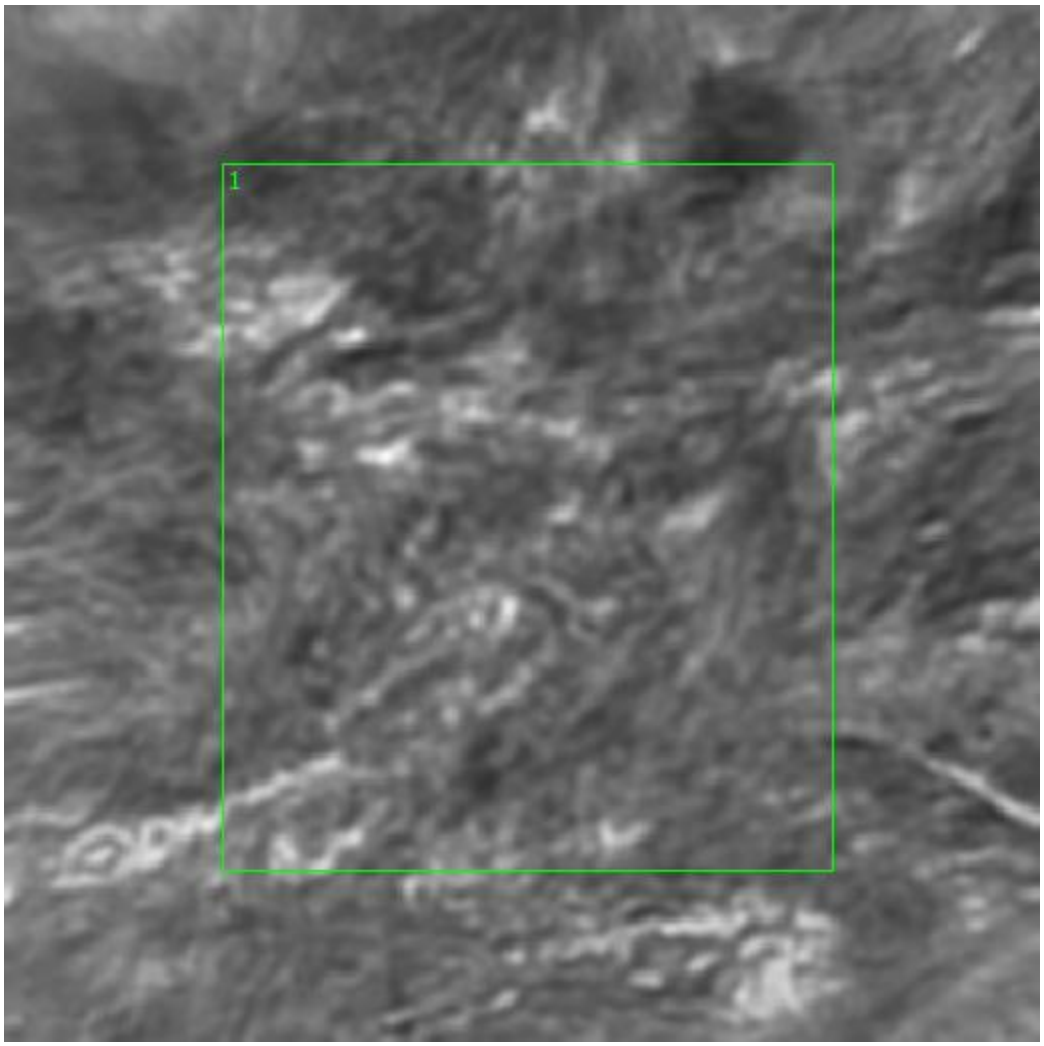



**Figure 0.78. Point L Regions of Interest**
**Region 1: basement membrane, lamina propia**

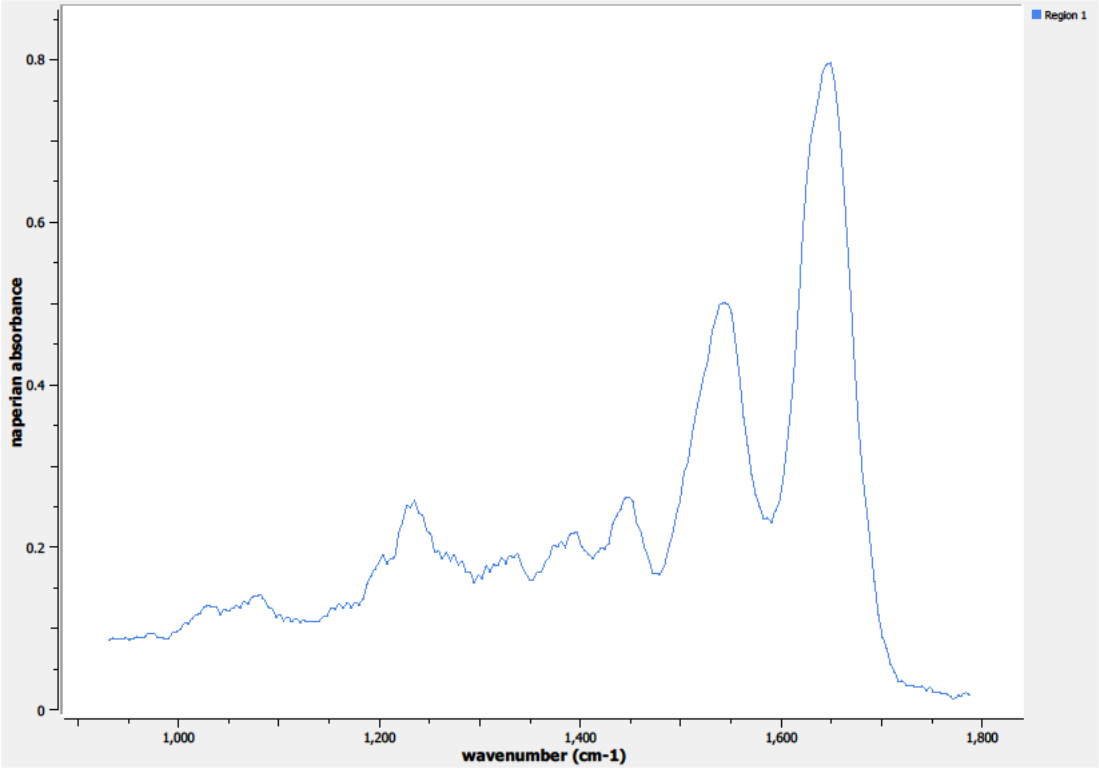



**Figure 0.79. Point L Spectra**

## A11. 12-05-A089a

| | |
|---|---|
| Unique ID | 12-05-A089a |
| Age | 42 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | no clearly-defined epithelium layer |

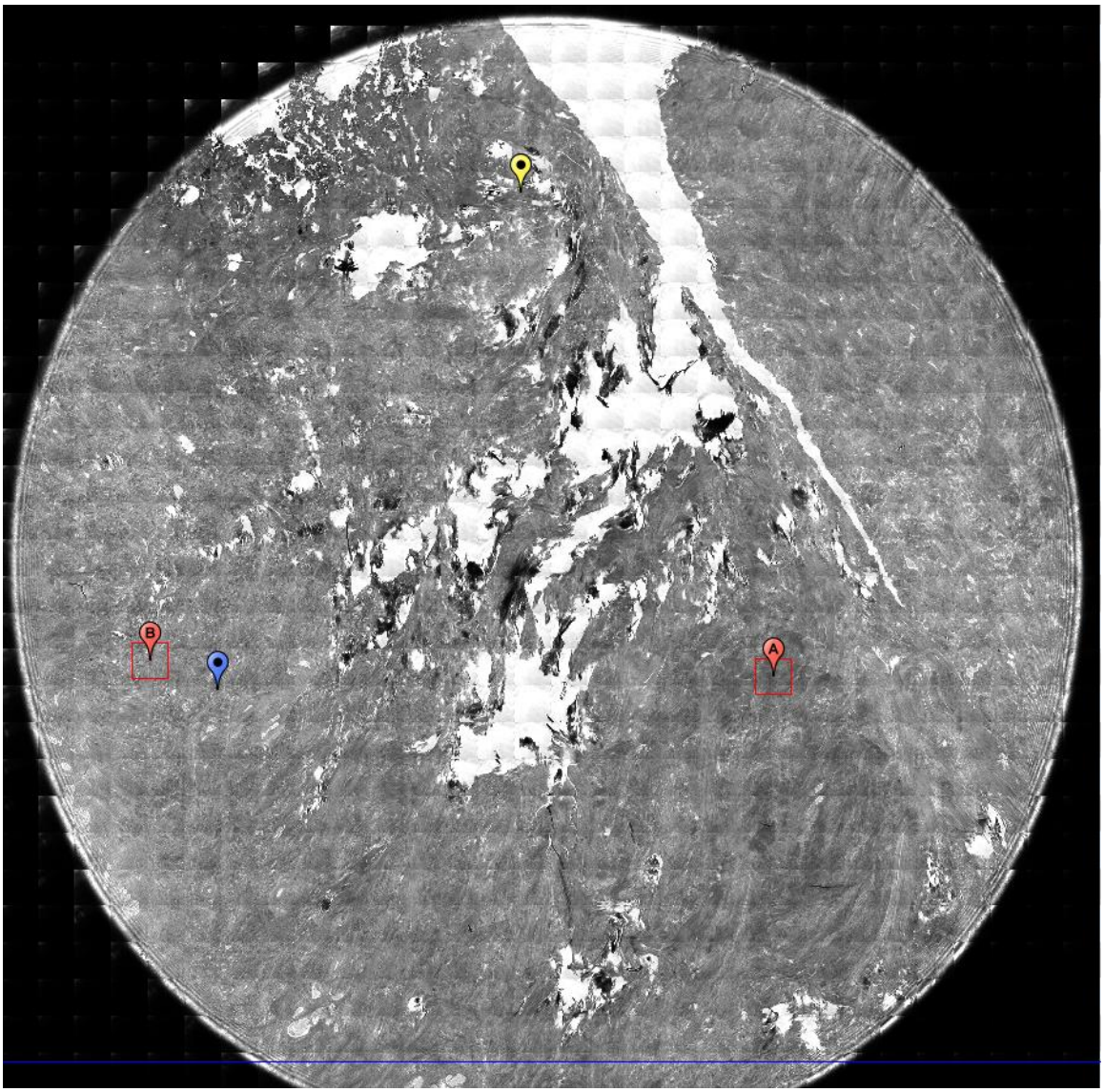



**Figure 0.80. Specimen 12-05-A089a, normal cervical tissue.**

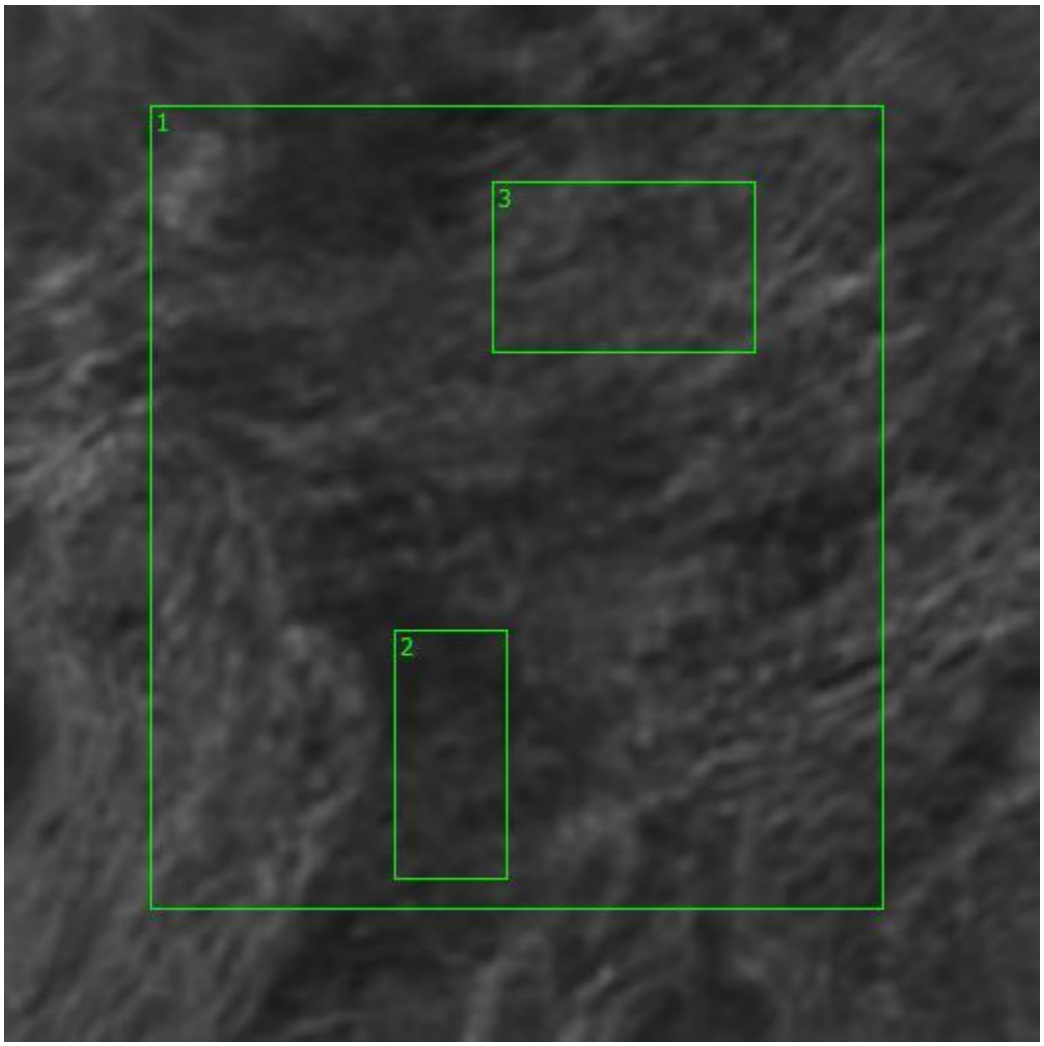



**Figure 0.81. Point A Regions of Interest**
**Regions 1, 2, 3: basement membrane, lamina propia**

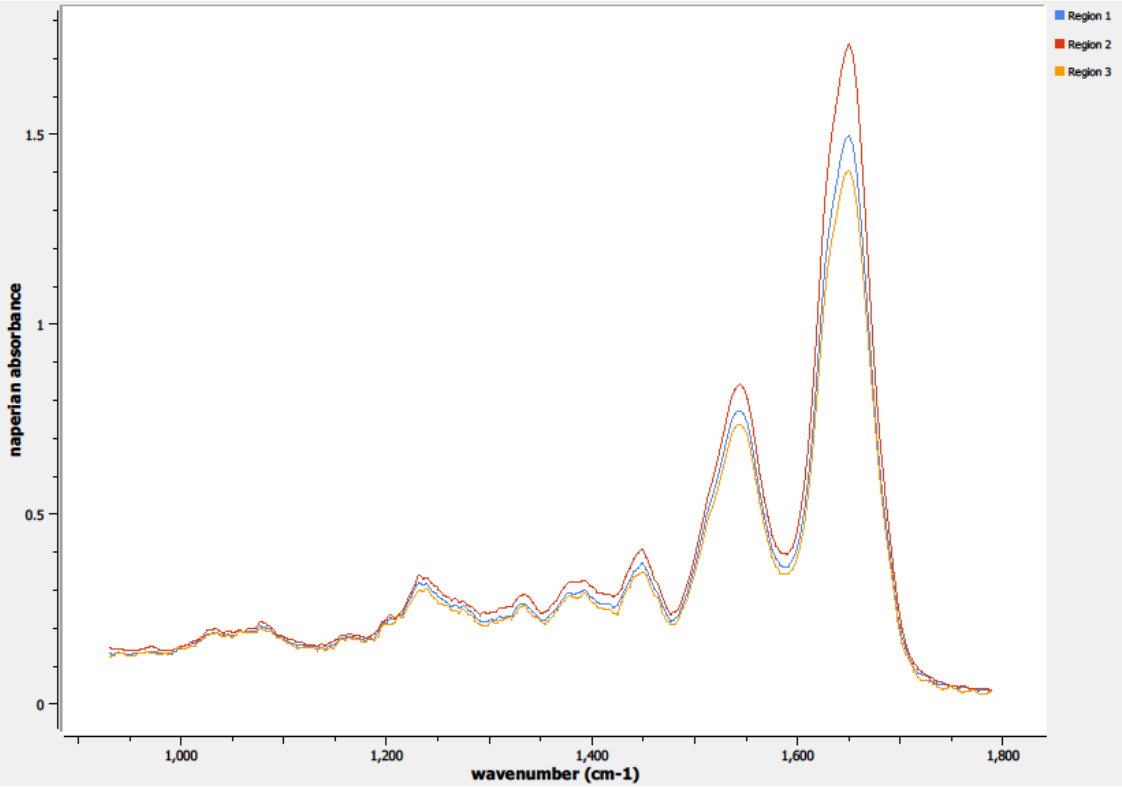



**Figure 0.82. Point A Spectra**

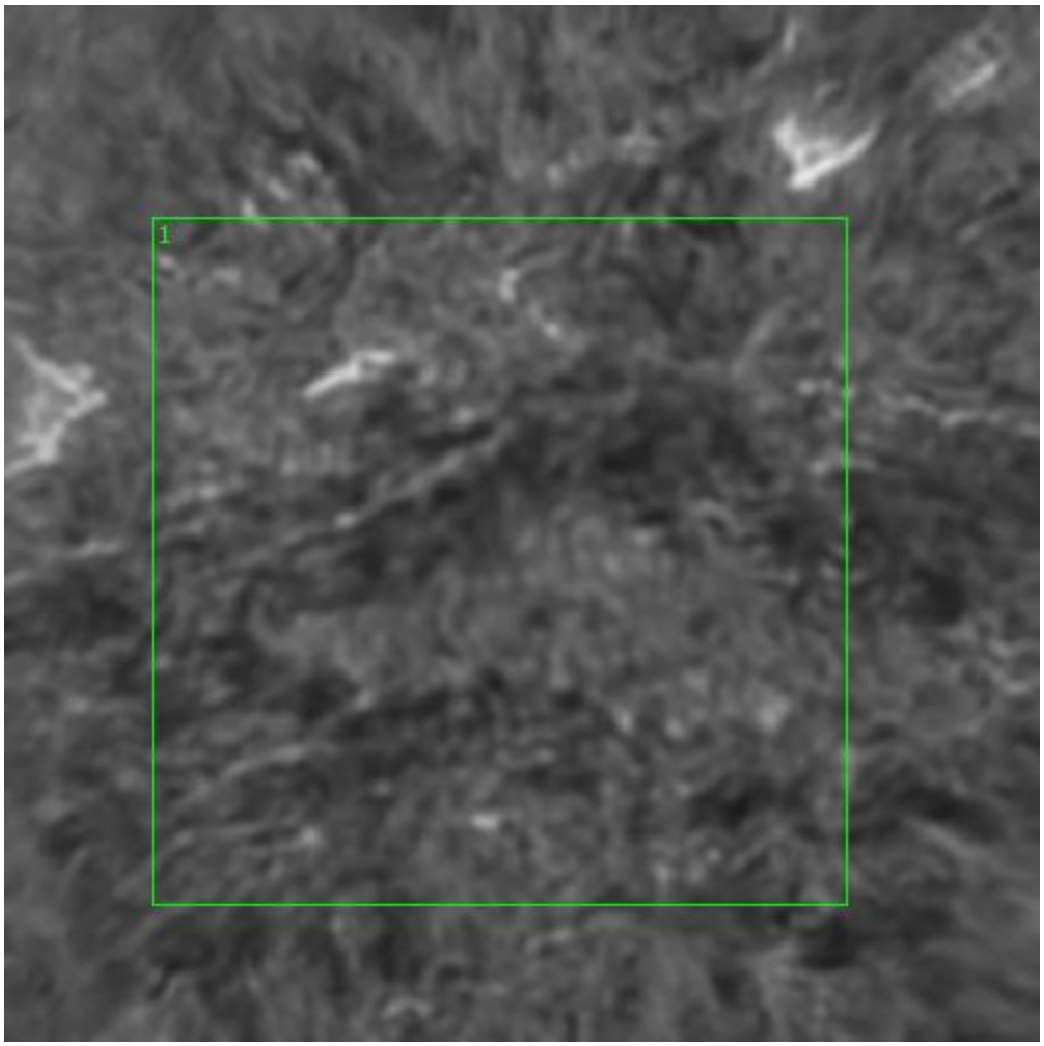



**Figure 0.83. Point B Regions of Interest**
**Region 1: basement membrane, lamina propia**

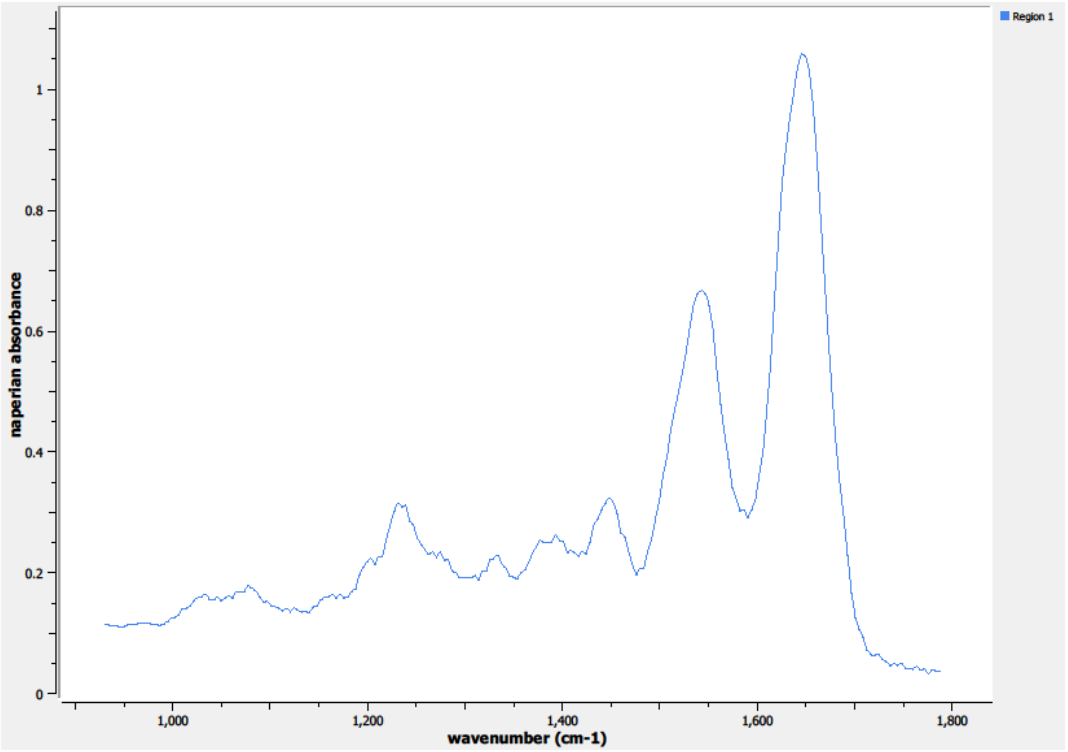



**Figure 0.84. Point B Spectra**

## A12. 12-05-A093a

| | |
|---|---|
| Unique ID | 12-05-A093a |
| Age | 54 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | none |

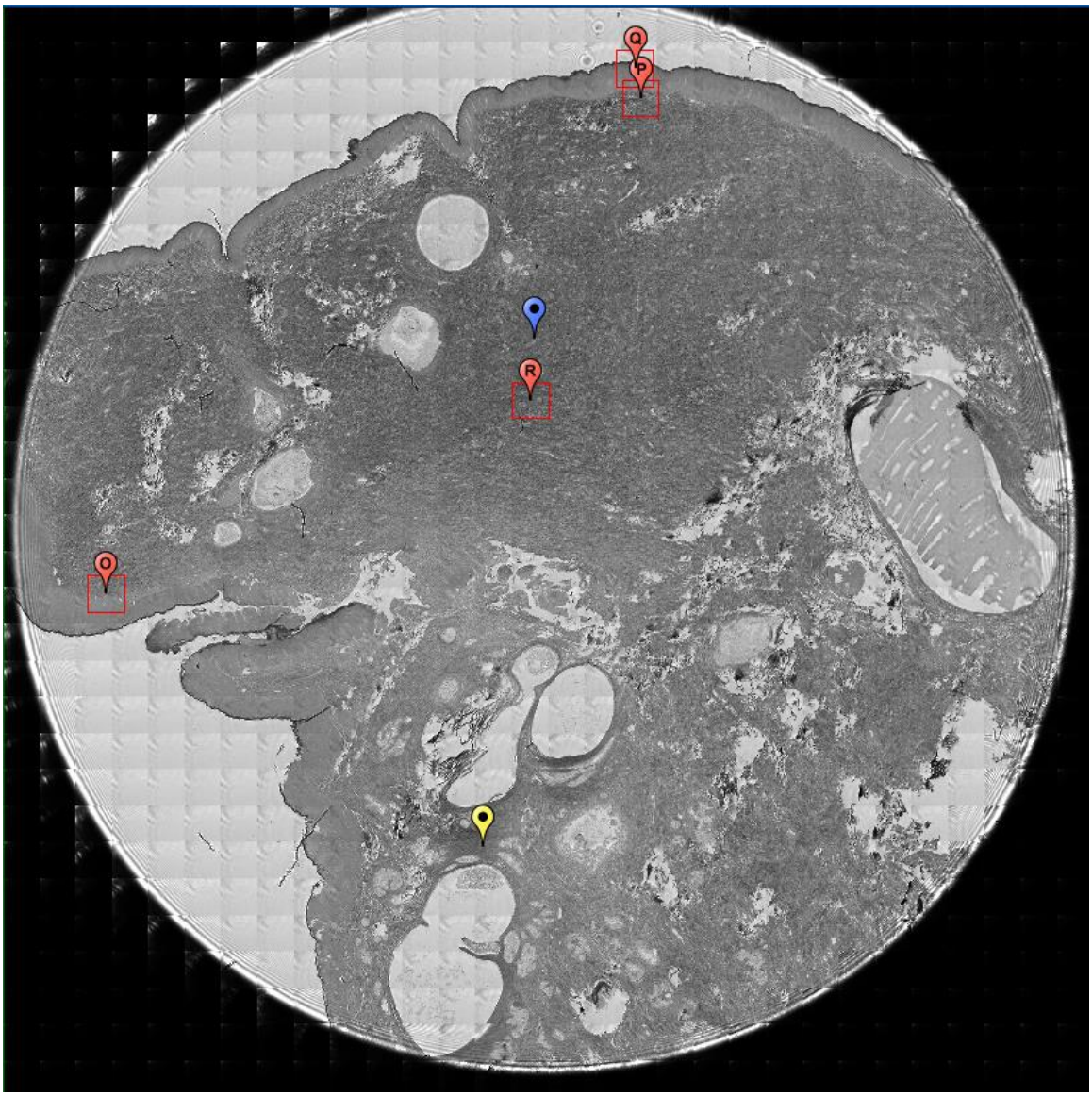



**Figure 0.85. Specimen 12-05-A093a, normal cervical tissue.**

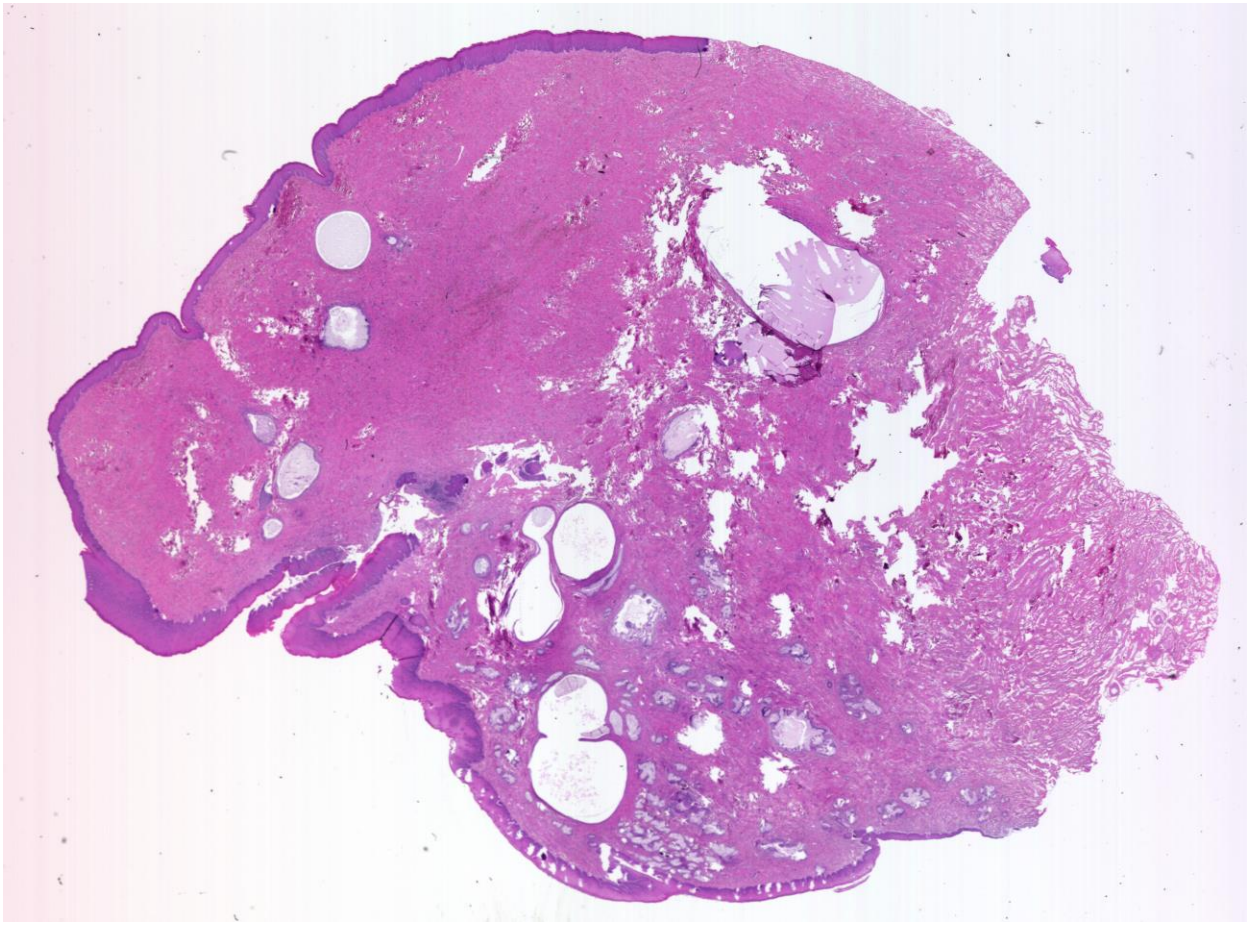


**Figure 0.86. Specimen 12-05-A093a, H&E stain, normal cervical tissue.**

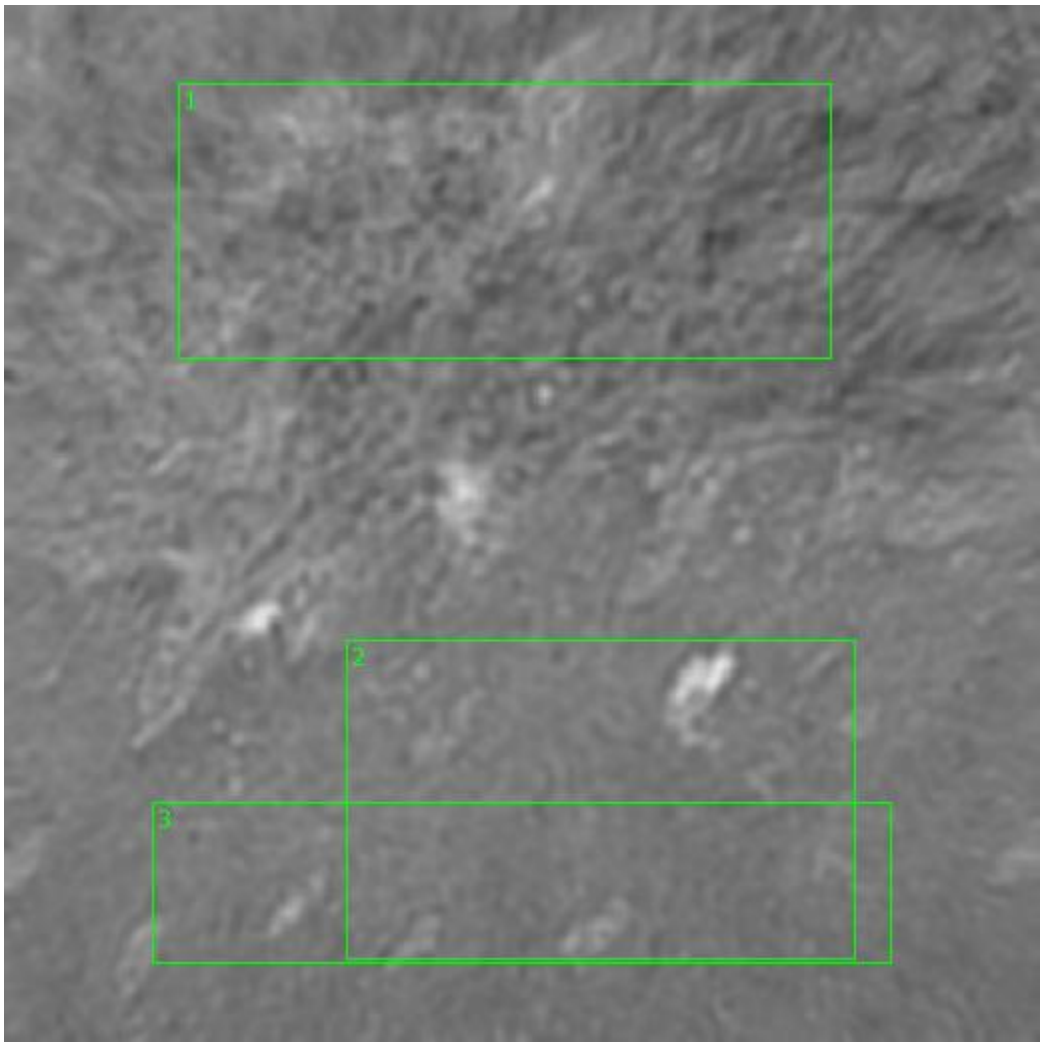



**Figure 0.87. Point O Regions of Interest**
**Region 1: basement membrane, lamina propia**
**Regions 2, 3: basal layer, parabasal layer**

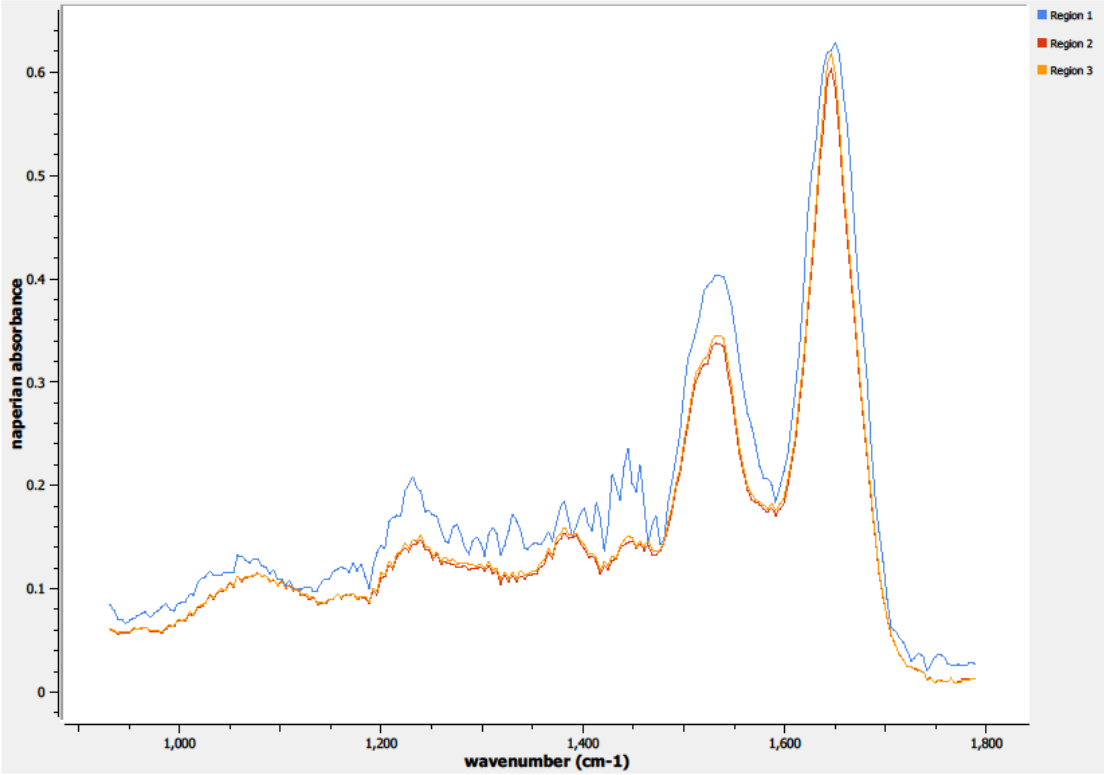



**Figure 0.88. Point O Spectra**

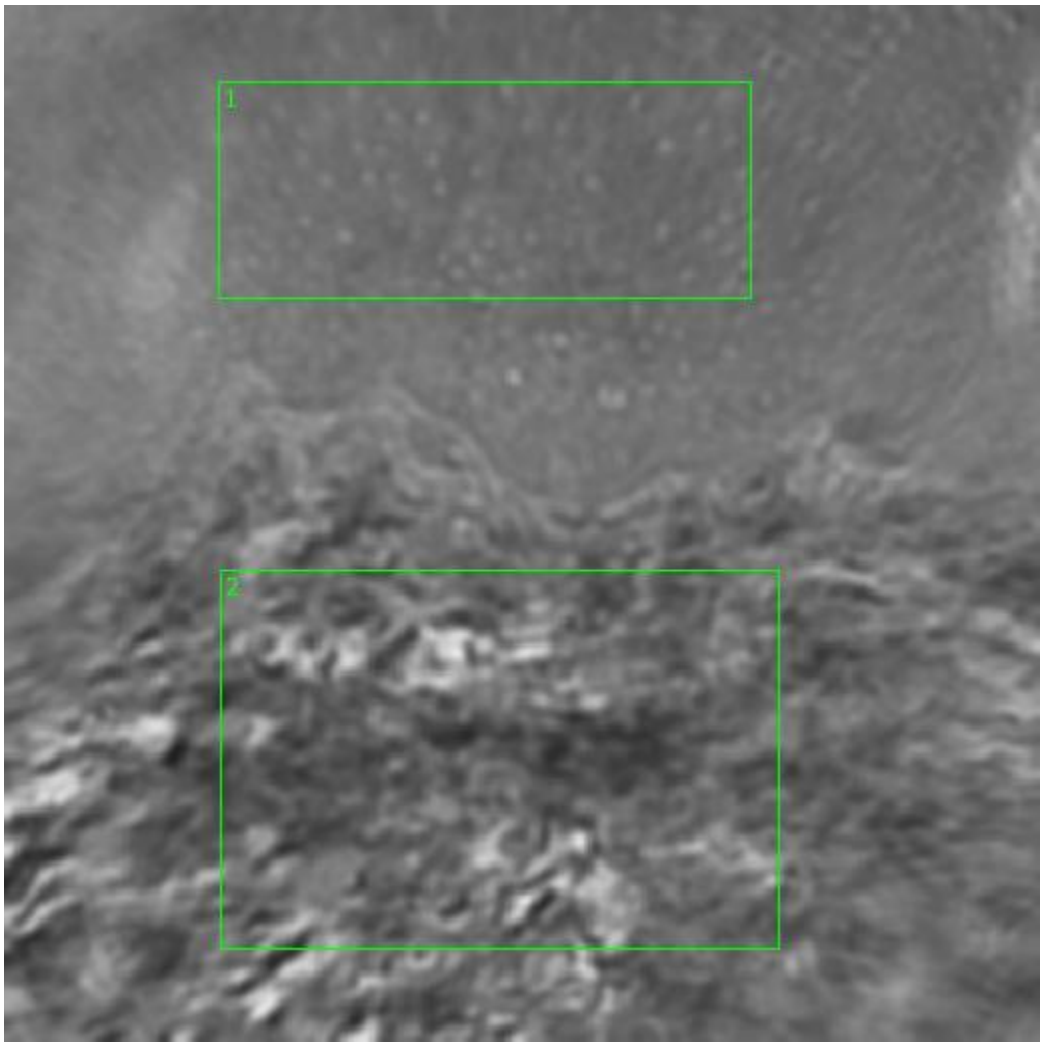



**Figure 0.89. Point P Regions of Interest**
**Region 1: parabasal layer**
**Region 2: basement membrane, lamina propia**

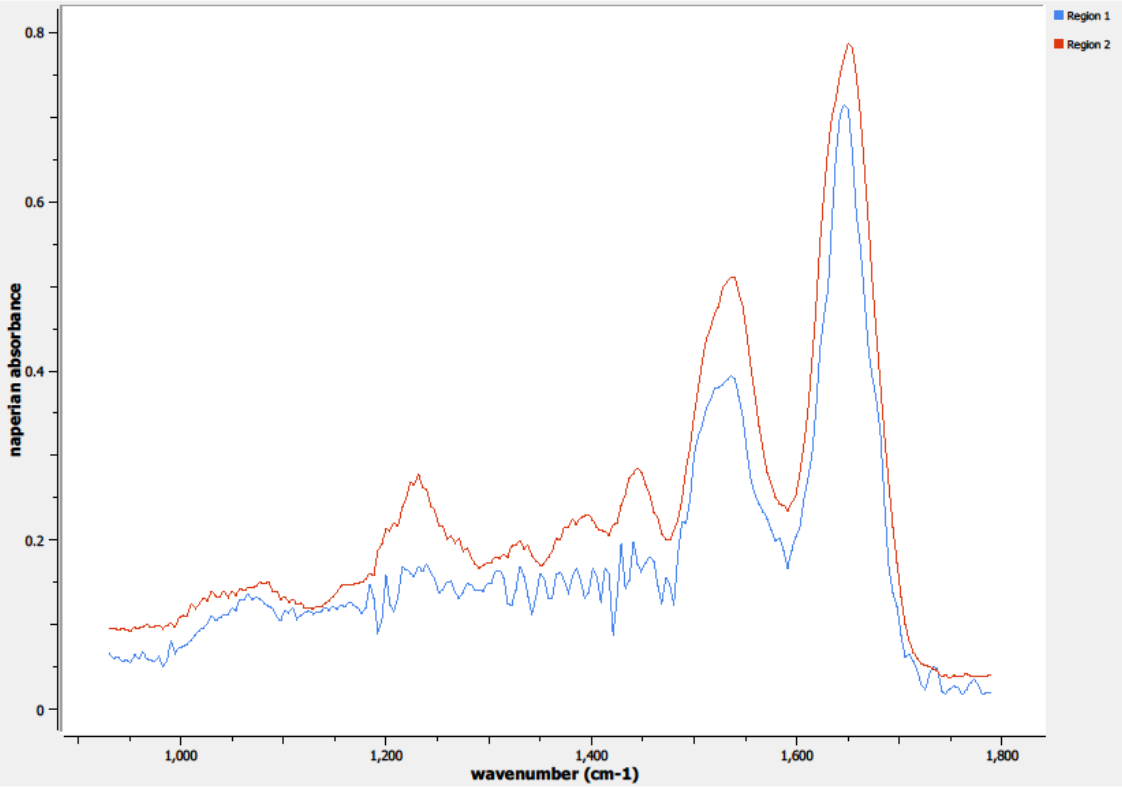



**Figure 0.90. Point P Spectra**

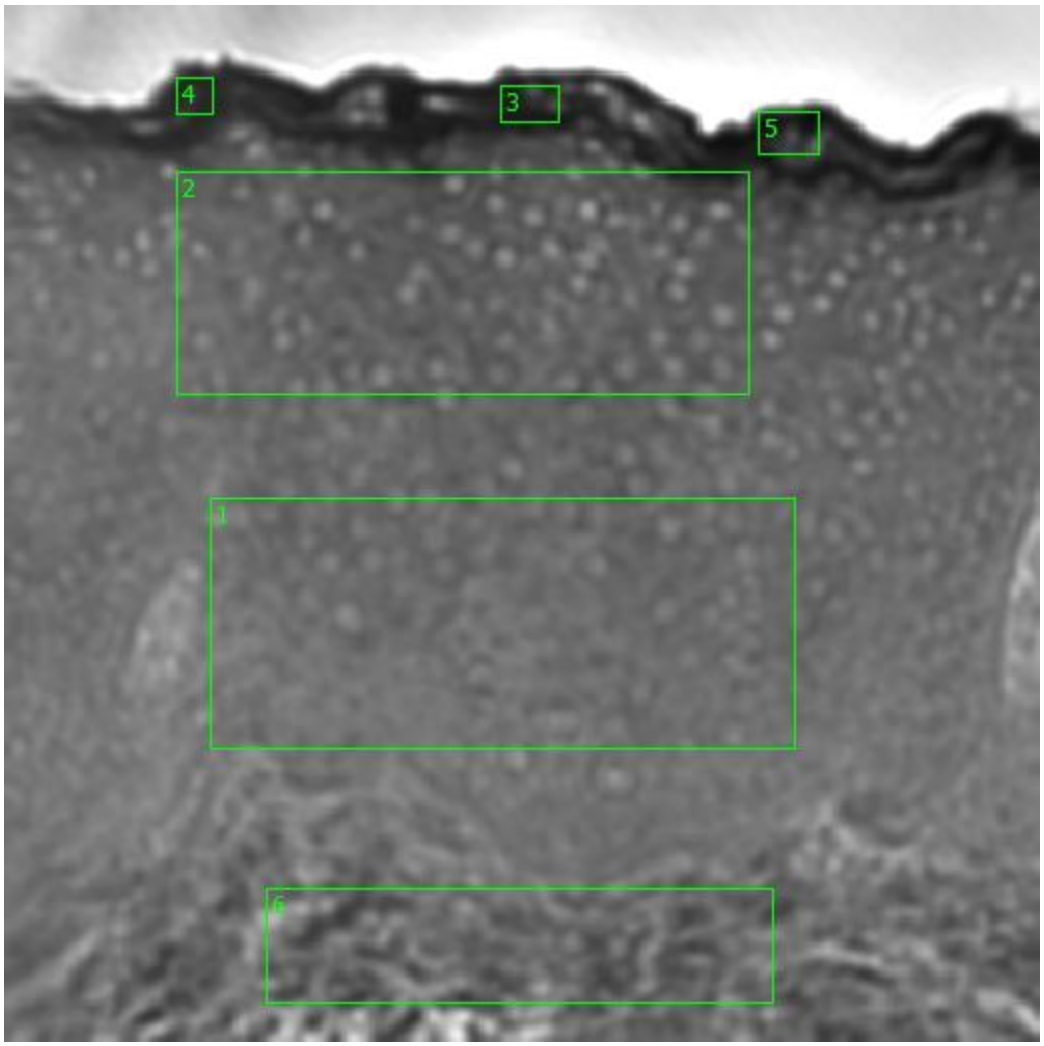



**Figure 0.91. Point Q Regions of Interest**
**Region 1: parabasal layer, columnar epithelium**
**Region 2: columnar epithelium**
**Regions 3, 4, 5: stratified squamous epithelium, epithelial surface**
**Region 6: basement membrane, lamina propia**

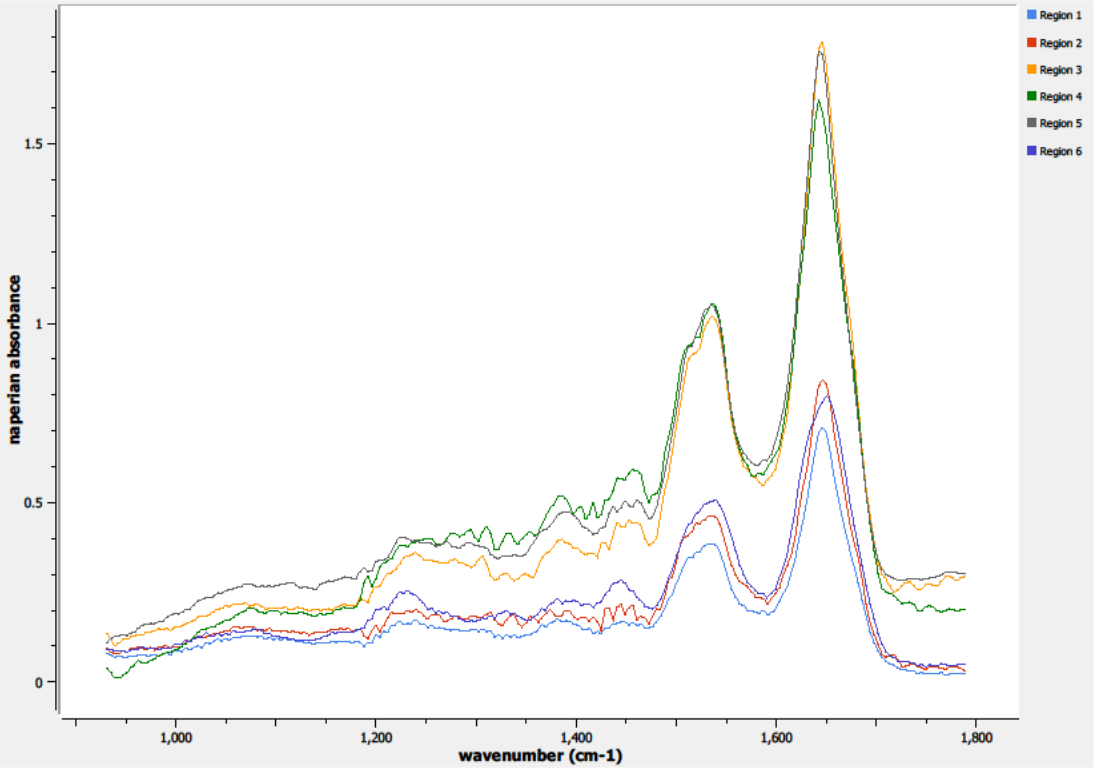



**Figure 0.92. Point Q Spectra**

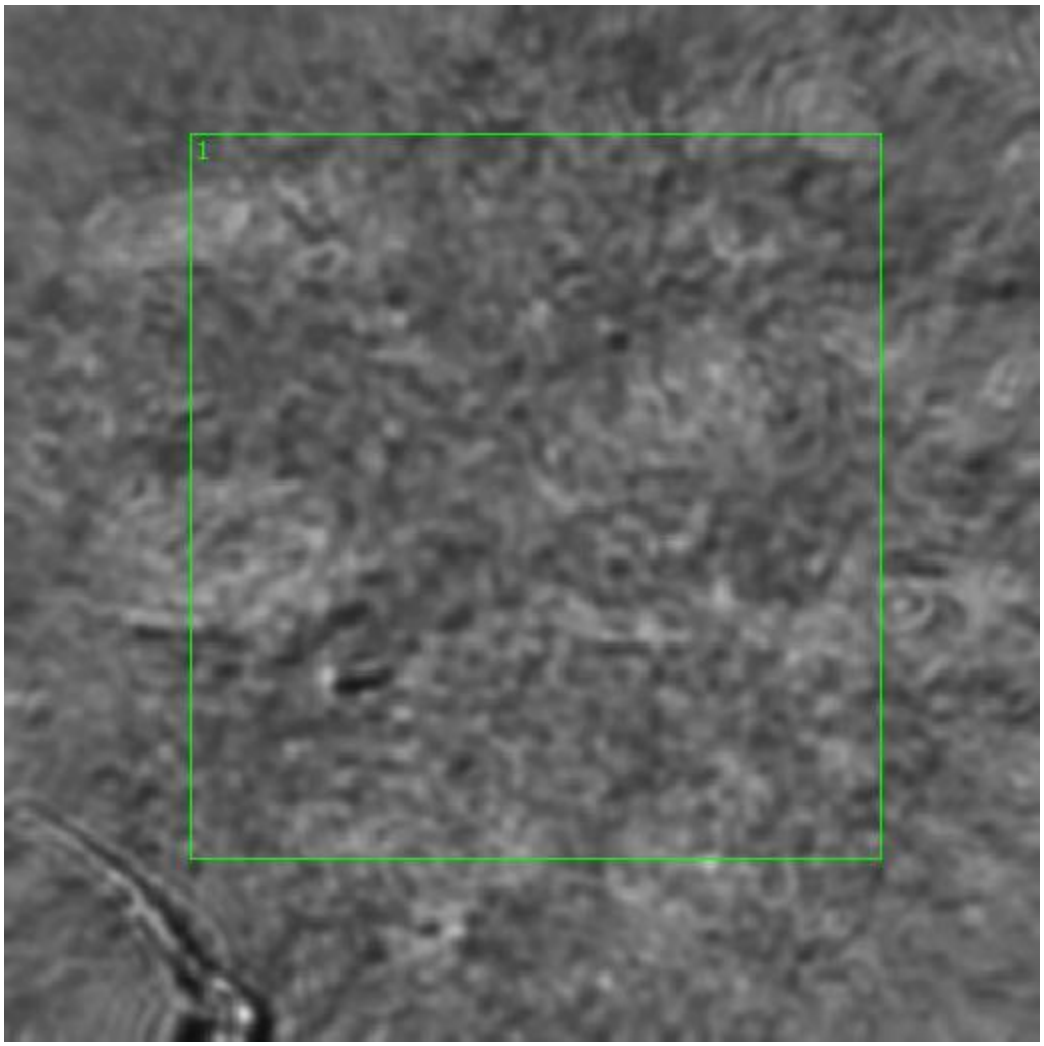



**Figure 0.93. Point R Regions of Interest**
**Region 1: lamina propia**

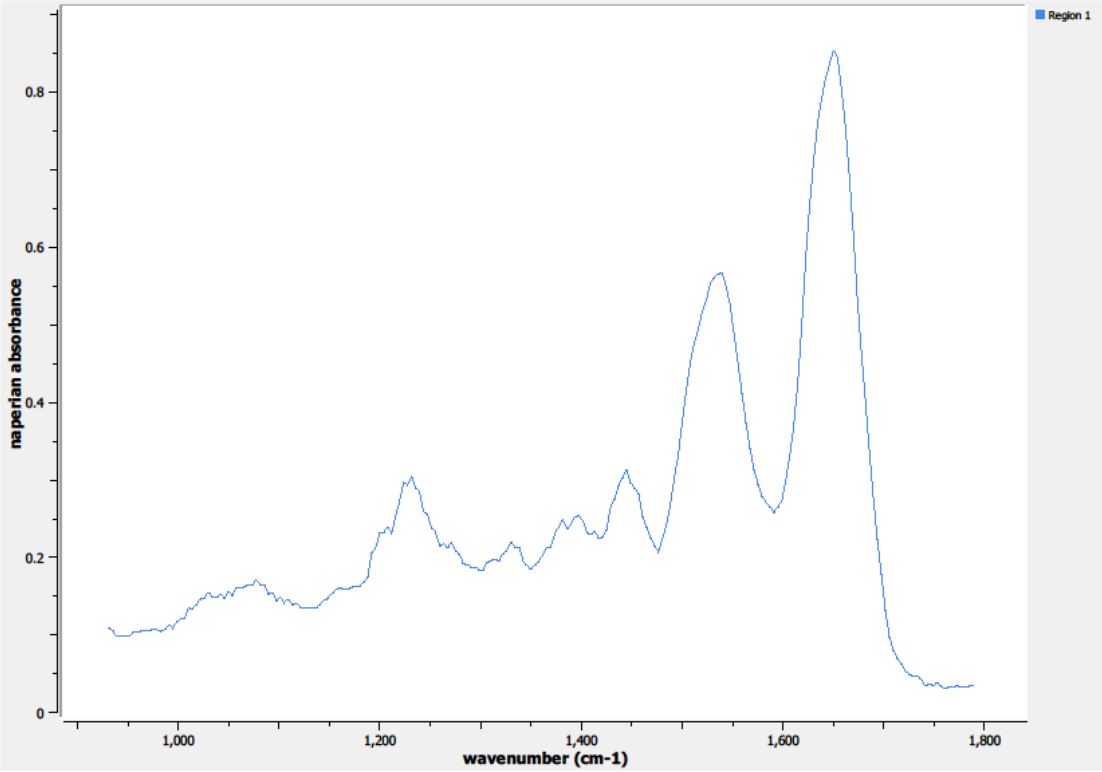



**Figure 0.94. Point R Spectra**

## A13. 12-05-A094a

| | |
|---|---|
| Unique ID | 12-05-A094a |
| Age | 34 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | no clearly-defined epithelium layer |

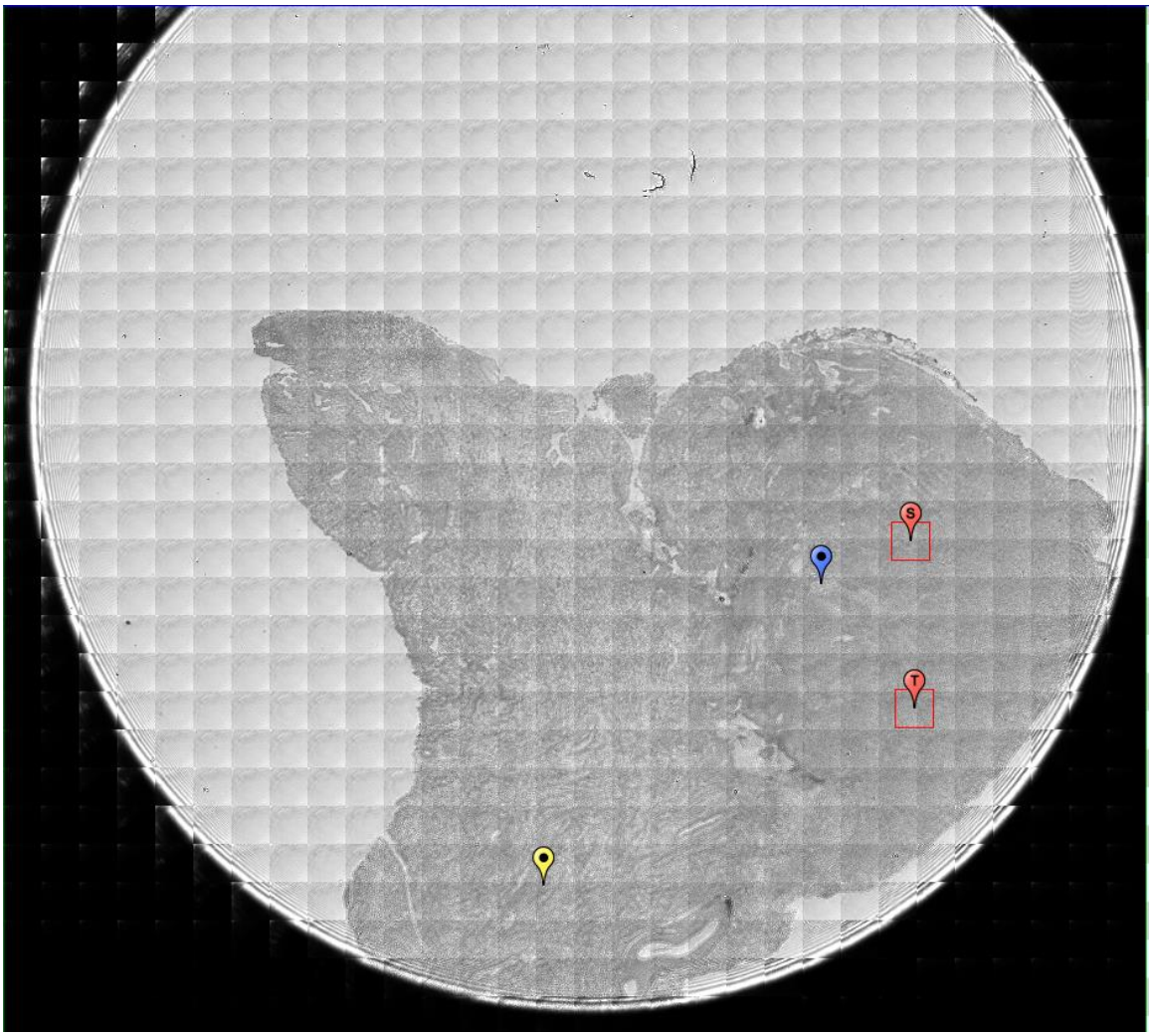



**Figure 0.95. Specimen 12-05-A094a, normal cervical tissue.**

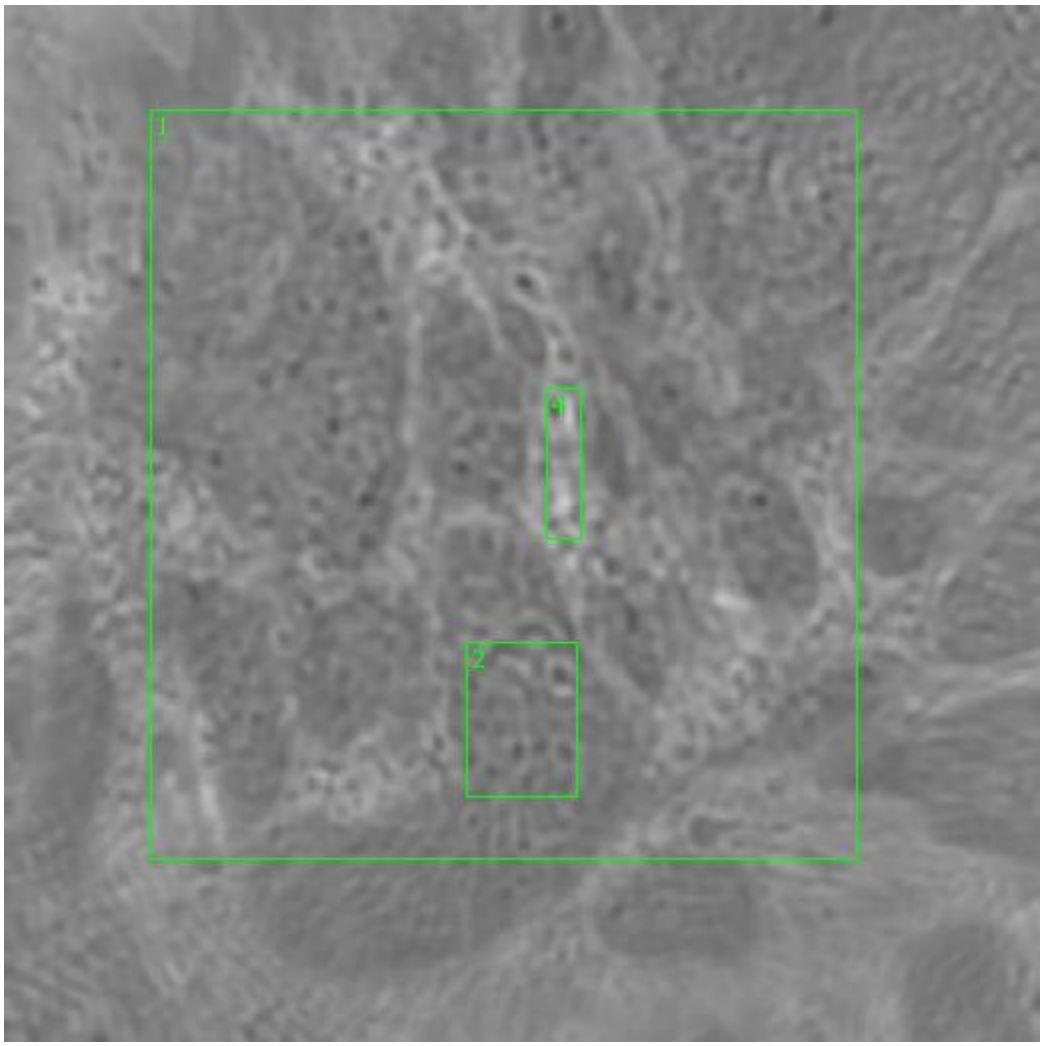



**Figure 0.96. Point S Regions of Interest**
**Regions 1, 2, 4: basement membrane, lamina propia**

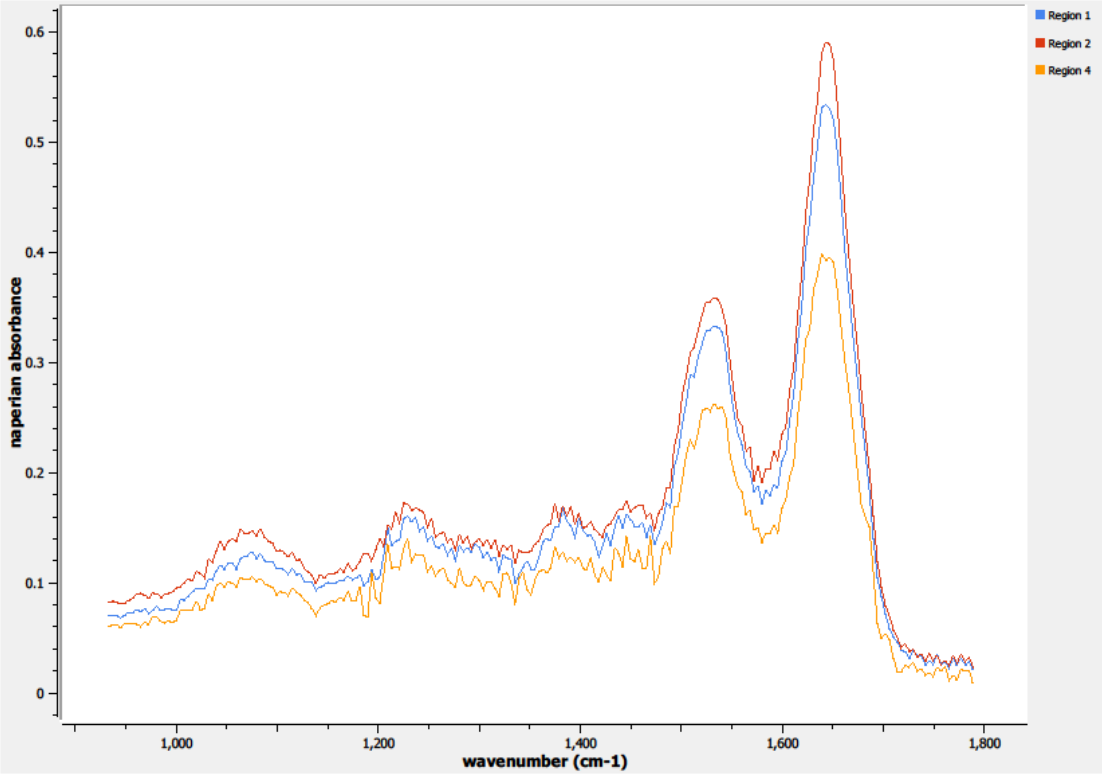



**Figure 0.97. Point S Spectra**

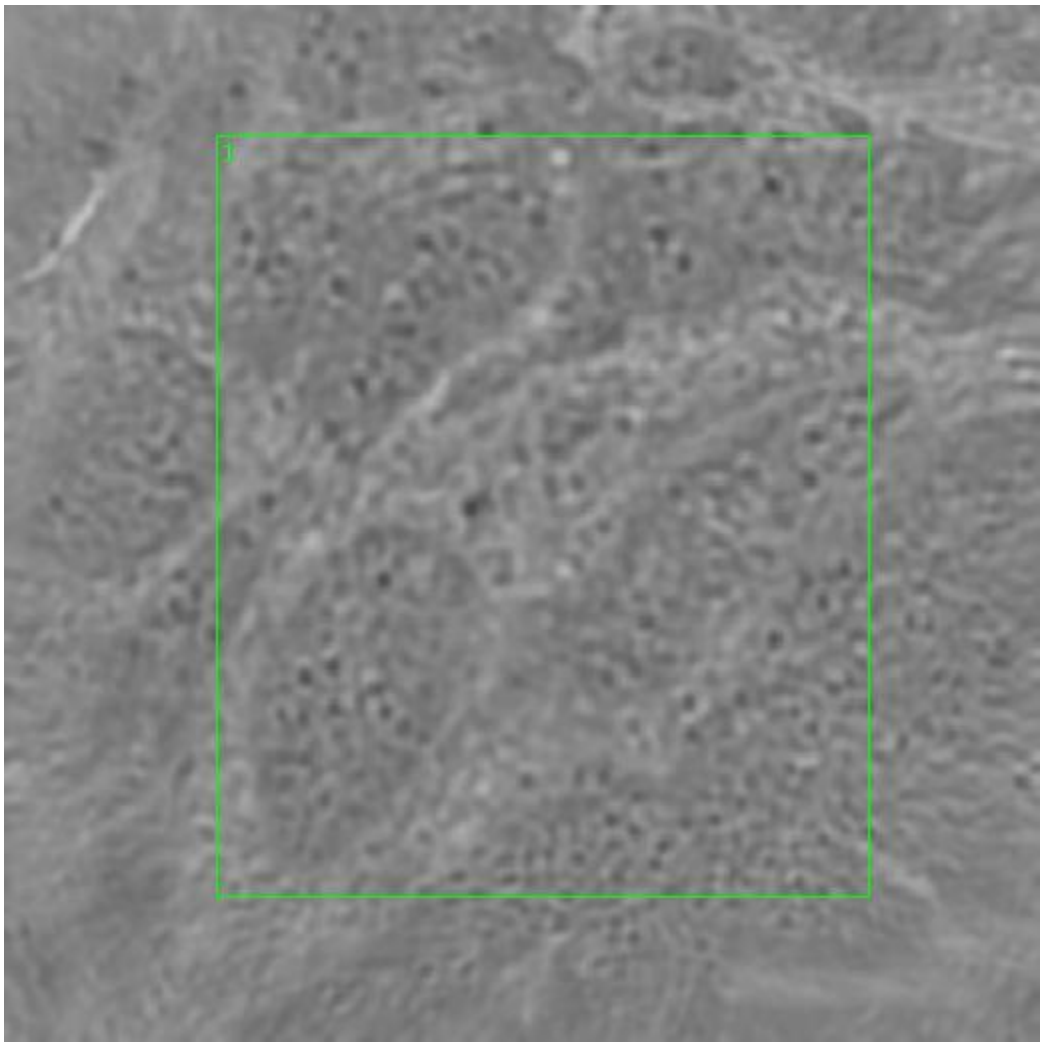



**Figure 0.98. Point T Regions of Interest**
**Region 1: basement membrane, lamina propia**

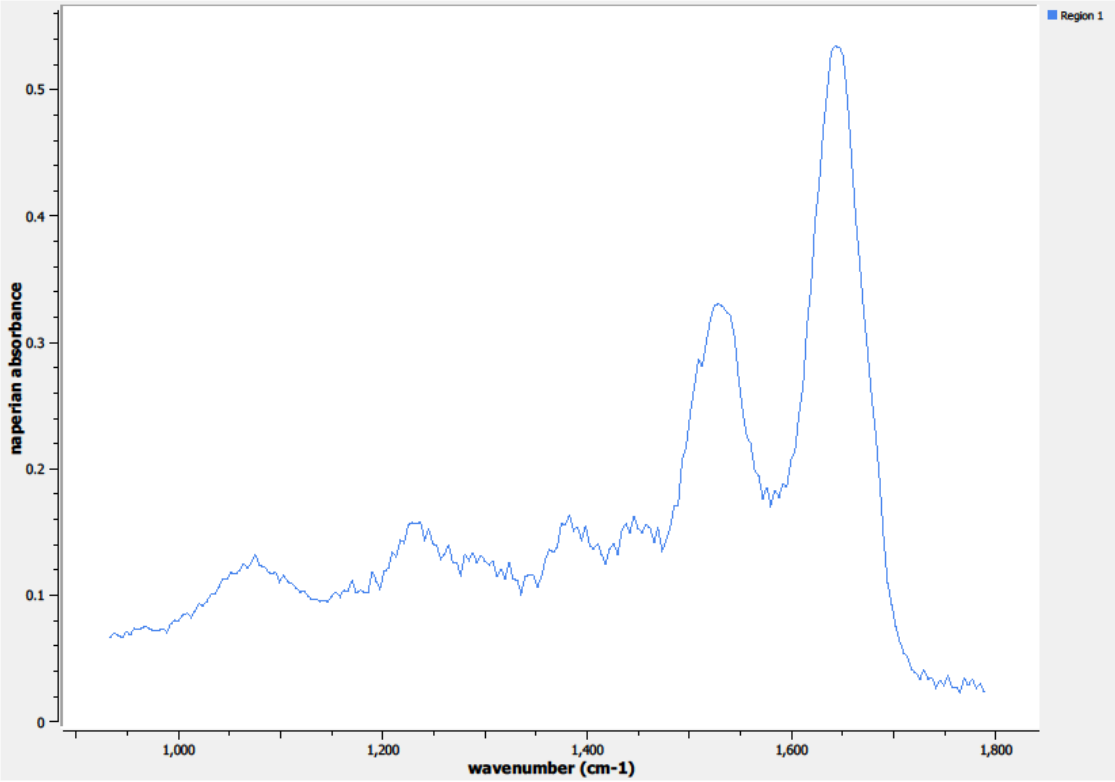



**Figure 0.99. Point T Spectra**

## A14. 12-05-A102a

| | |
|---|---|
| Unique ID | 12-05-A102a |
| Age | 53 |
| Race | white |
| Location of Sample Collection | endocervix |
| Notes of Interest | none |

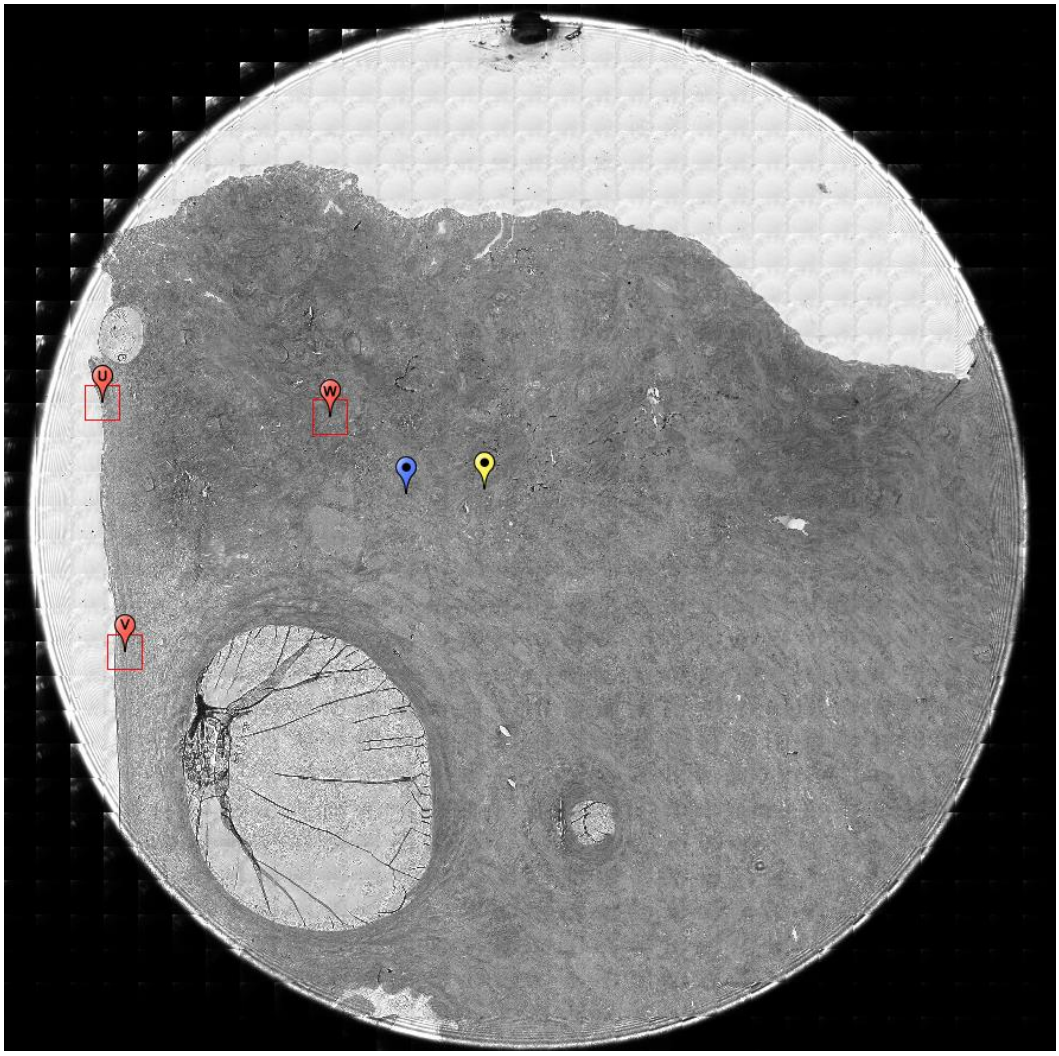



**Figure 0.100. Specimen 12-05-A102a, normal cervical tissue.**

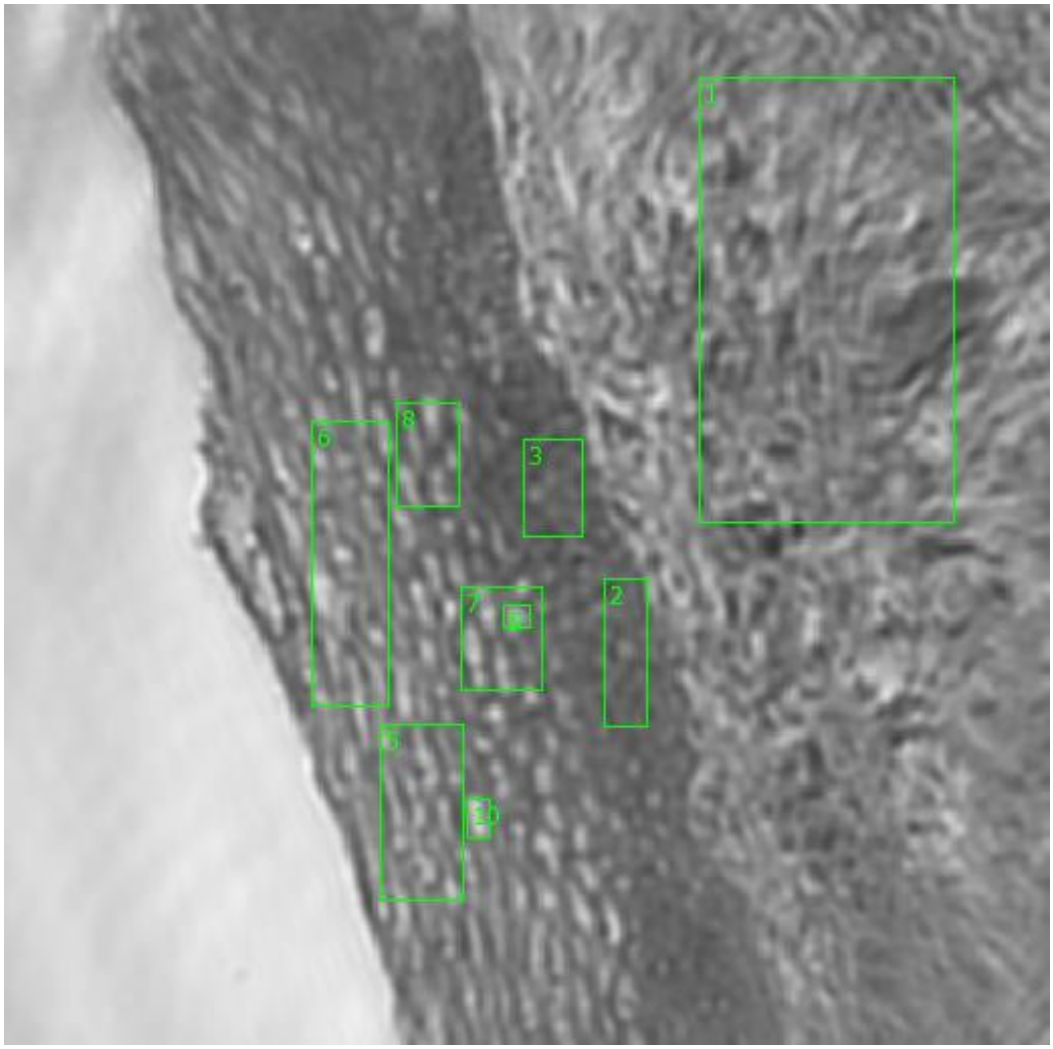



**Figure 0.101. Point U Regions of Interest**
**Region 1: basement membrane**
**Regions 2, 3: basal layer**
**Regions 5, 6, 7, 8: stratified squamous epithelium**
**Regions 9, 10: individual stratified squamous epithelial cells**

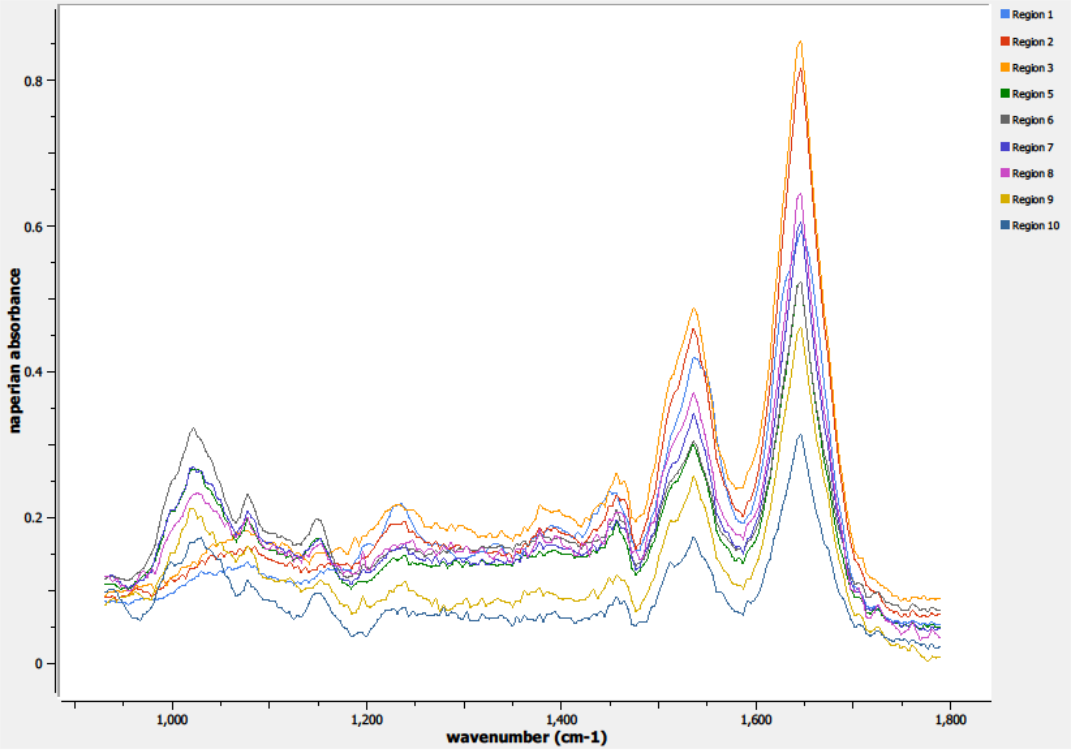



**Figure 0.102. Point U Spectra**

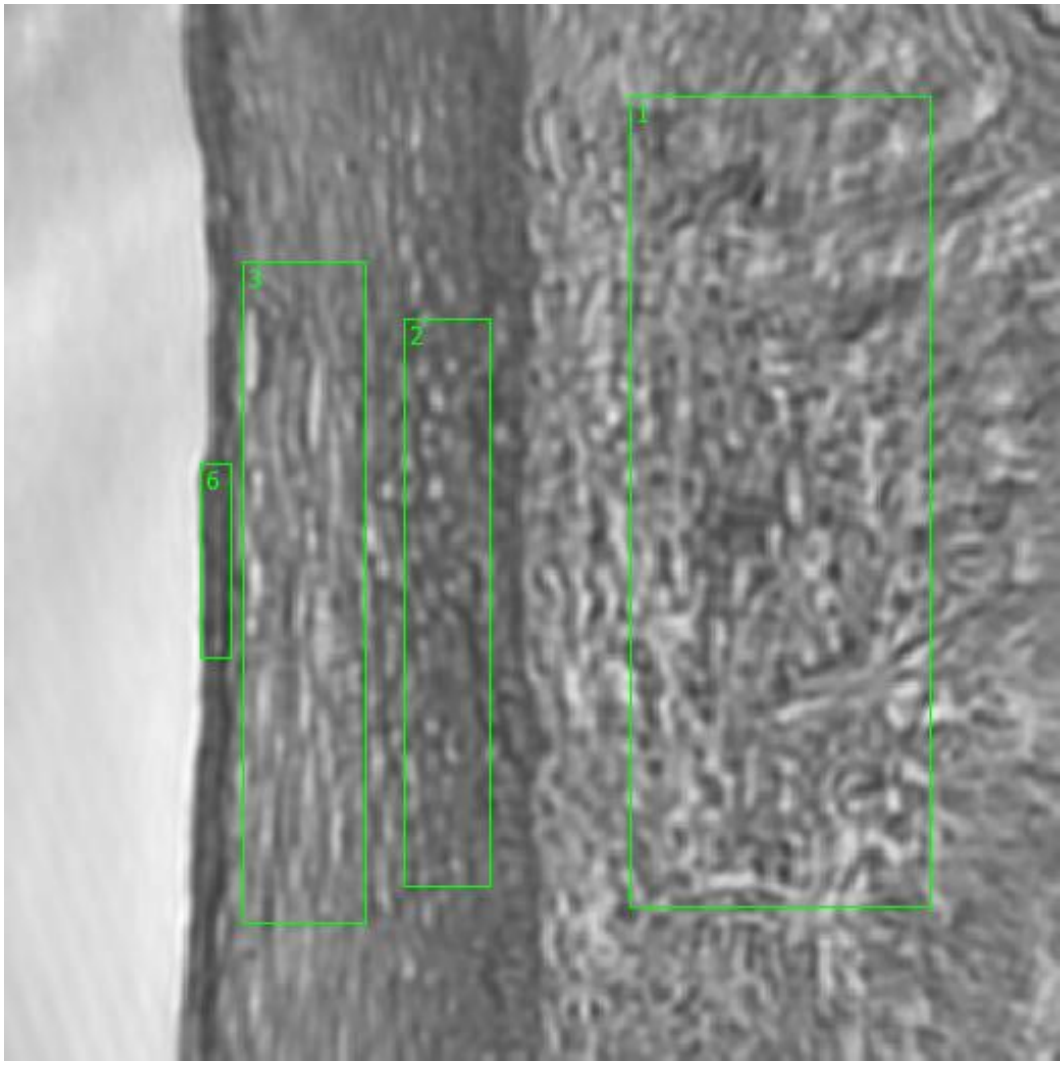



**Figure 0.103. Point V Regions of Interest**
**Region 1: basement membrane**
**Region 2: columnar epithelium, parabasal layer**
**Region 3: stratified squamous epithelium, intermediate**
**Region 6: epithelial surface**

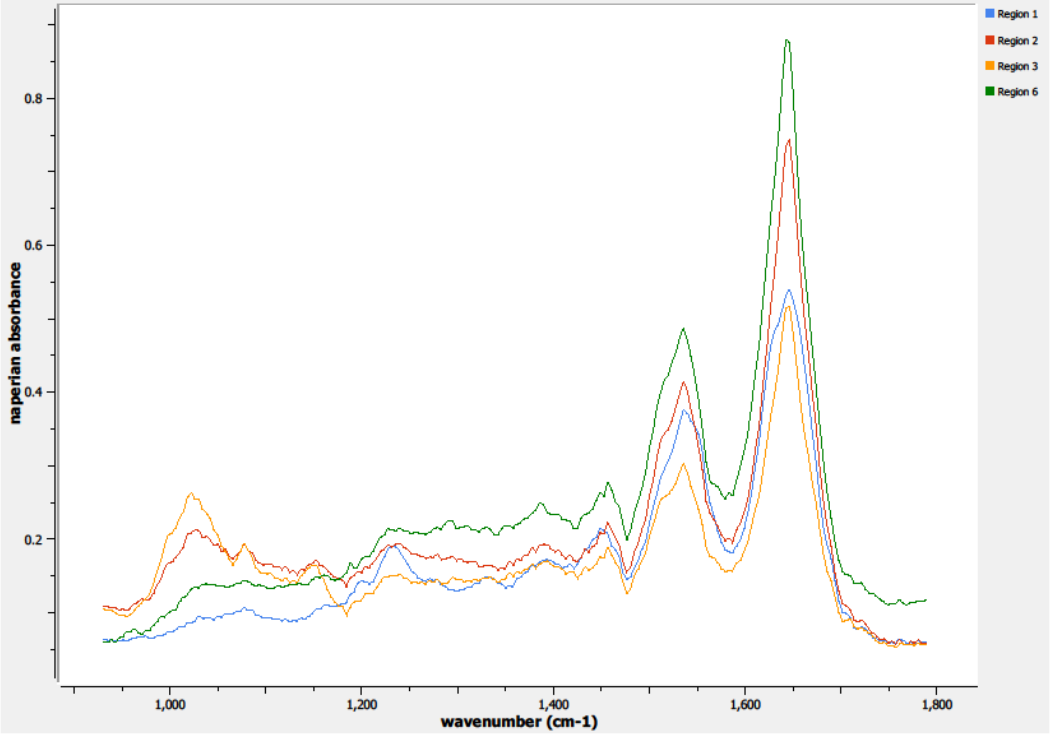



**Figure 0.104. Point V Spectra**

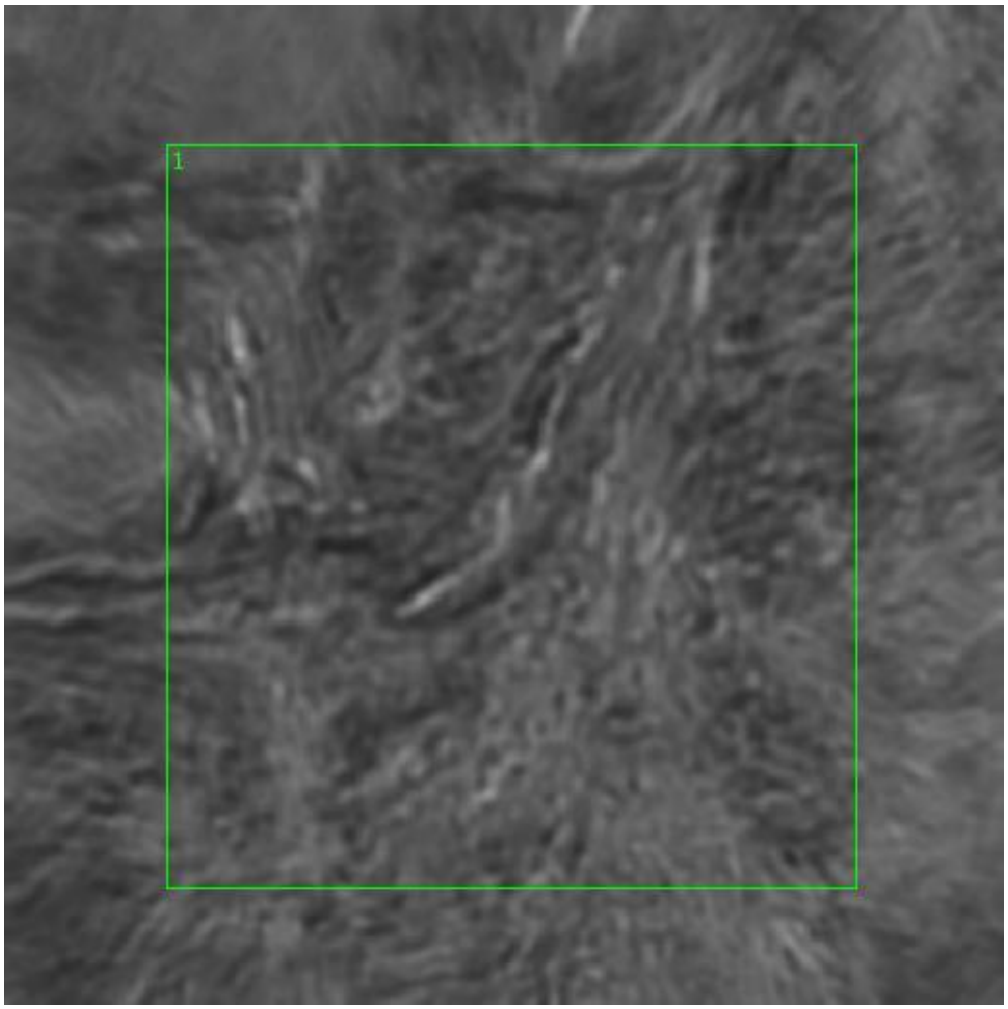



**Figure 0.105. Point W Regions of Interest**
**Region 1: lamina propia**

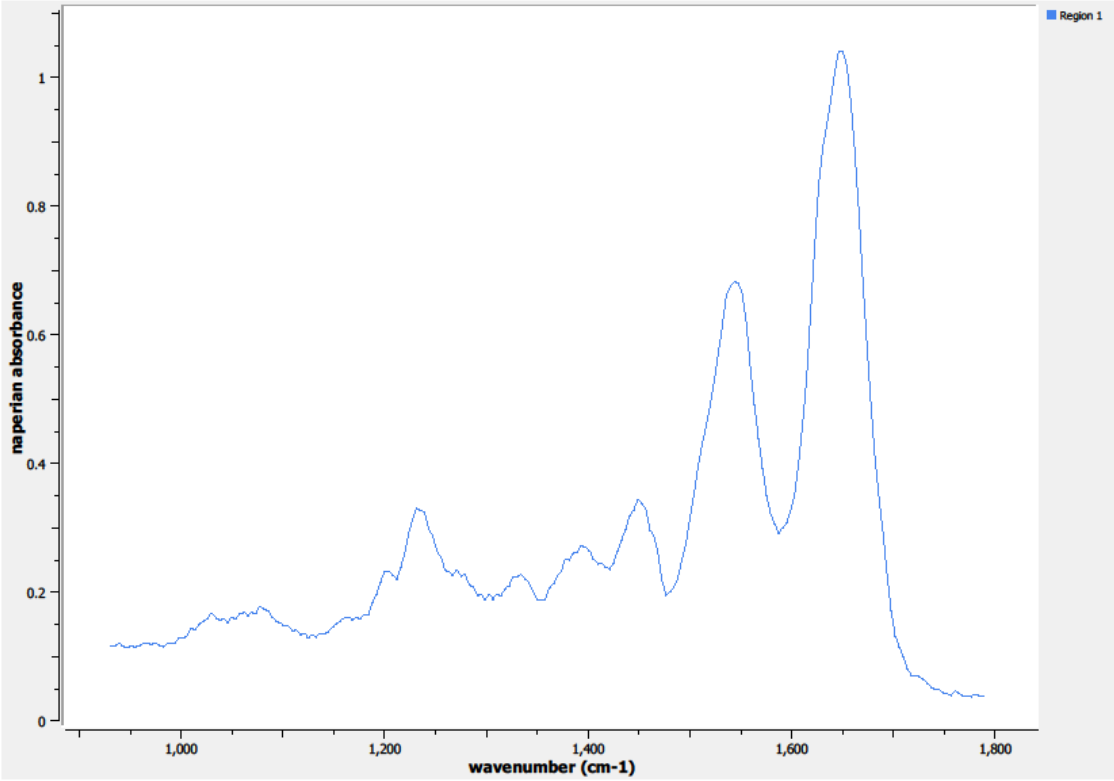



**Figure 0.106. Point W Spectra**

## A15. 12-05-A131a

| | |
|---|---|
| Unique ID | 12-05-A131a |
| Age | 42 |
| Race | black |
| Location of Sample Collection | ectocervix |
| Notes of Interest | none |

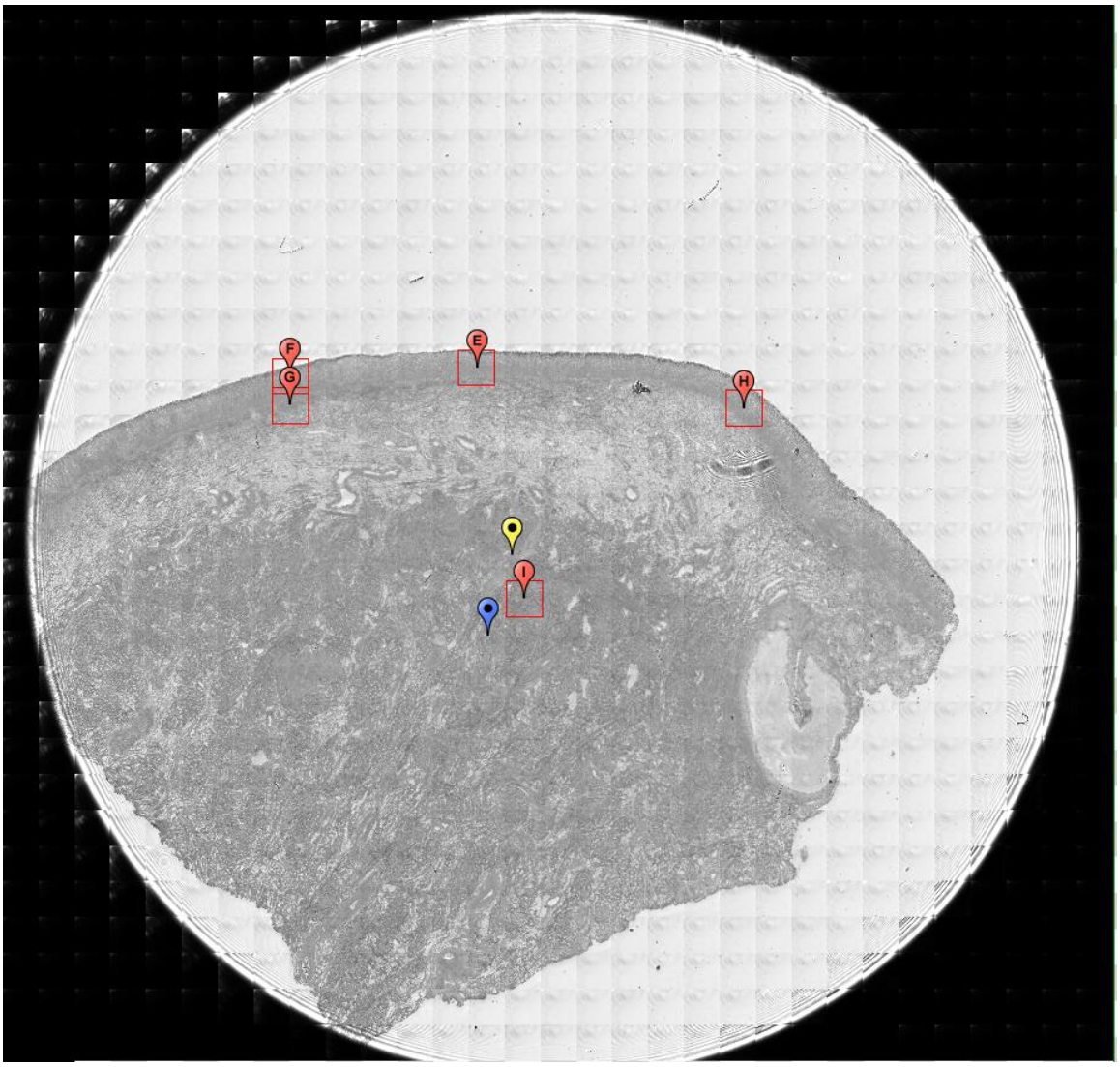



**Figure 0.107. Specimen 12-05-A131a, normal cervical tissue.**

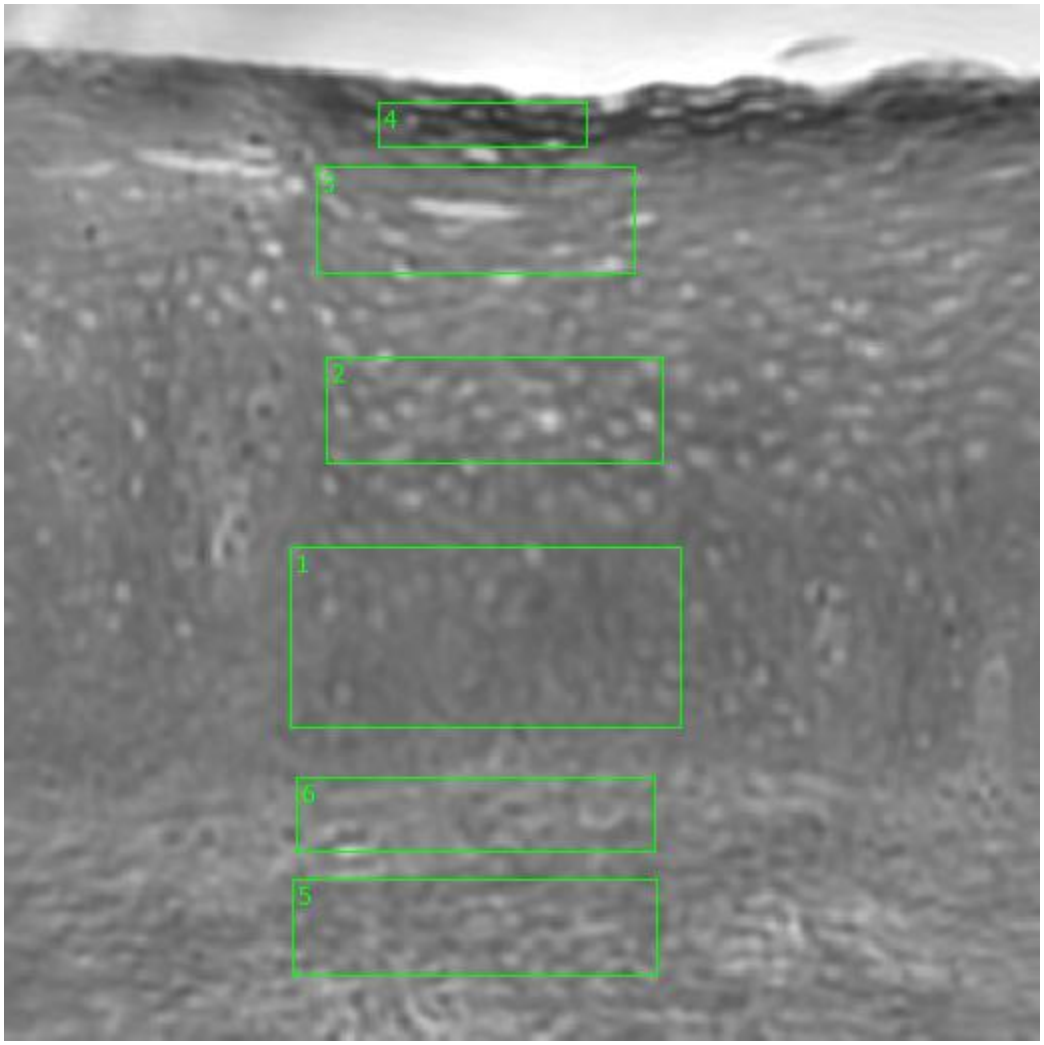



**Figure 0.108. Point E Regions of Interest**
**Region 1: parabasal layer**
**Region 2: columnar epithelium**
**Region 3: stratified squamous epithelium**
**Region 4: stratified squamous epithelium, epithelial surface**
**Regions 5, 6: basement membrane**

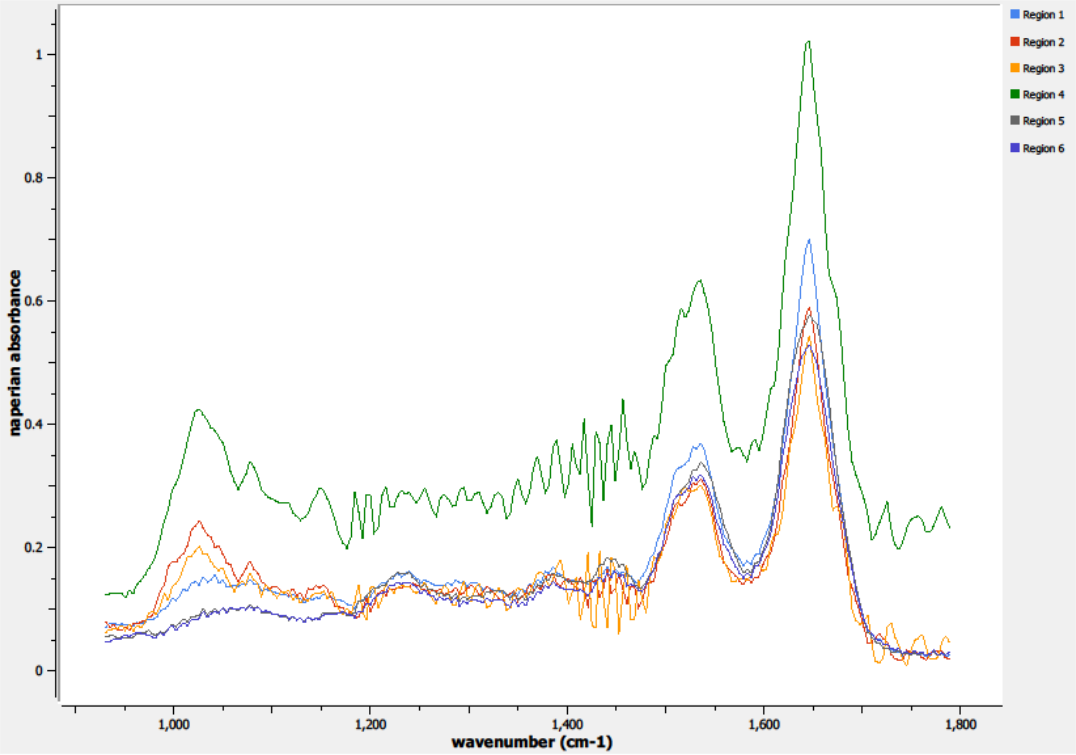



**Figure 0.109. Point E Spectra**

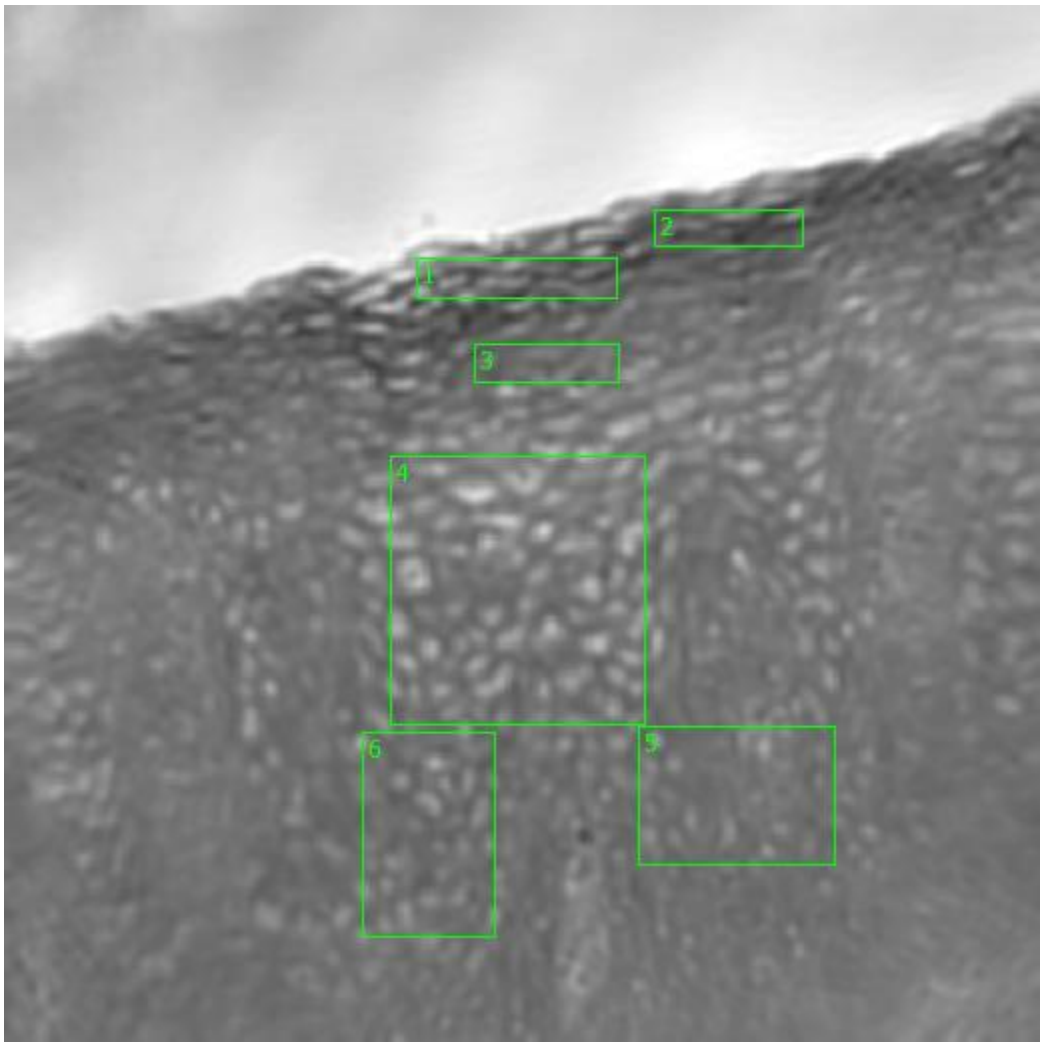



**Figure 0.110. Point F Regions of Interest**
**Regions 1, 2: stratified squamous epithelium, epithelial surface**
**Region 3: stratified squamous epithelium**
**Region 4: columnar epithelium**
**Regions 5, 6: parabasal layer**

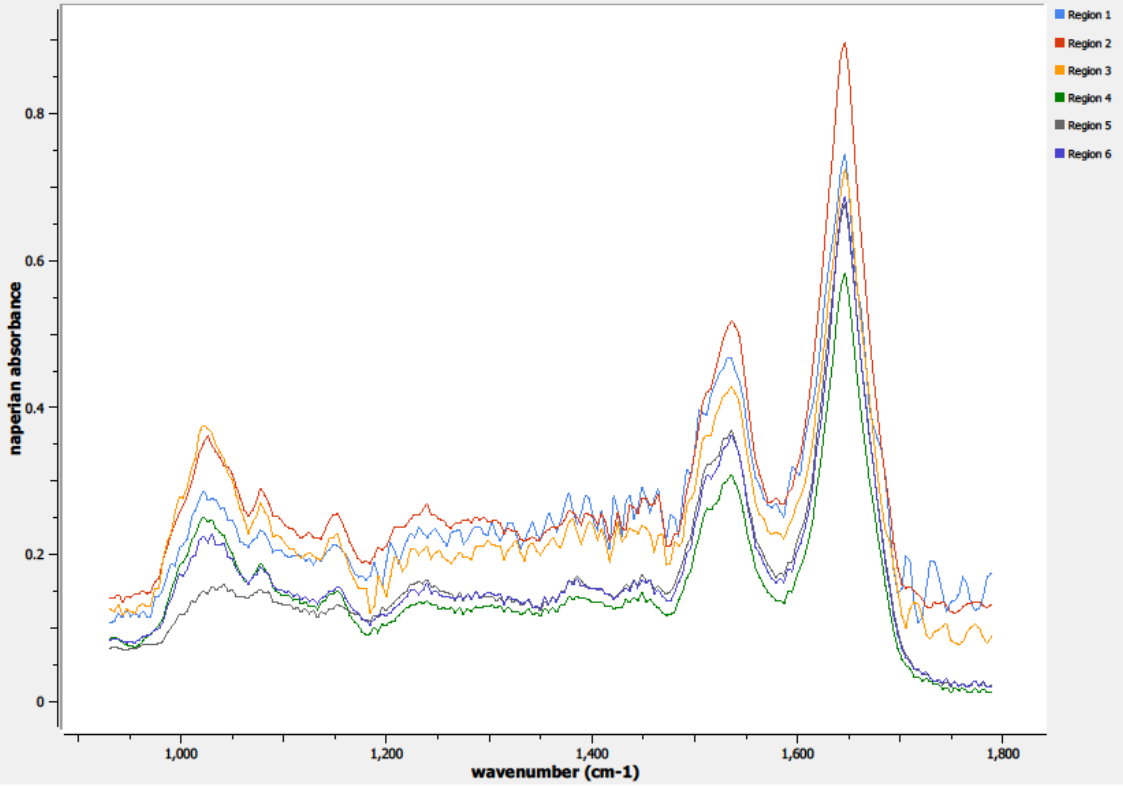



**Figure 0.111. Point F Spectra**

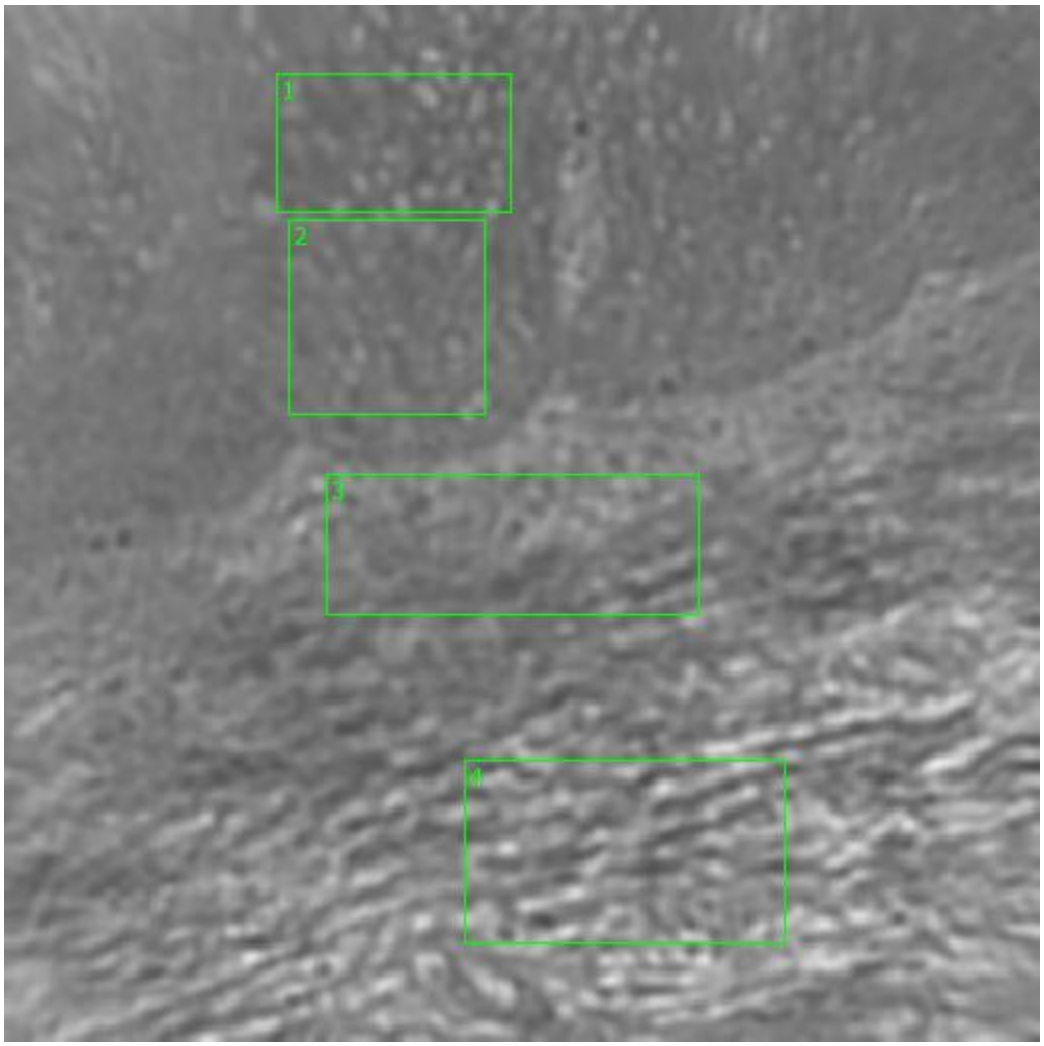



**Figure 0.112. Point G Regions of Interest**
**Regions 1, 2: parabasal layer**
**Regions 3, 4: basement membrane**

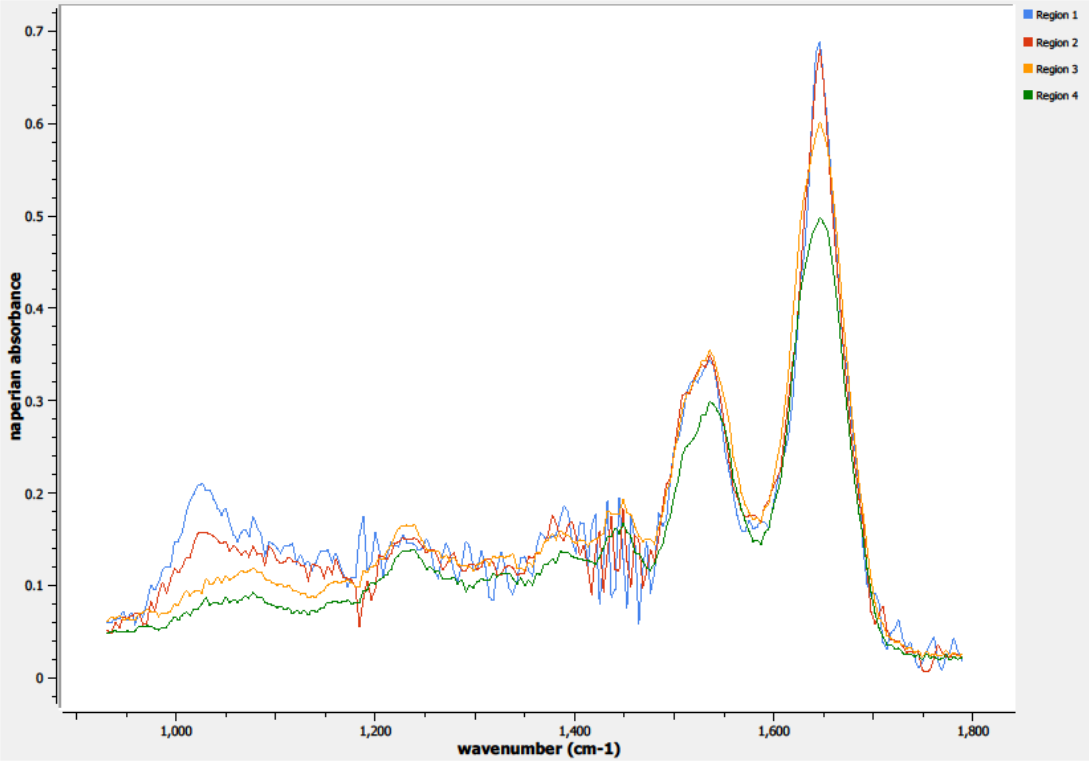



**Figure 0.113. Point G Spectra**

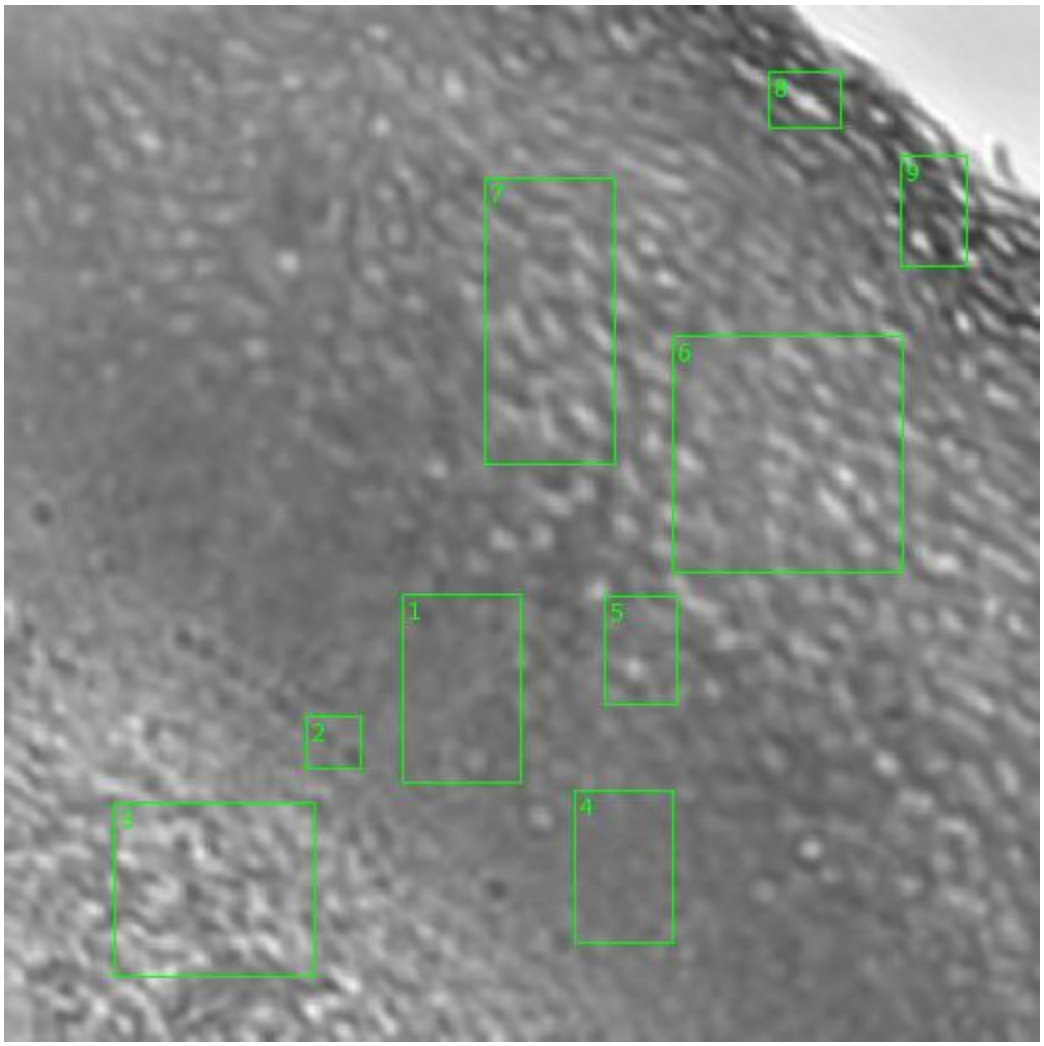



**Figure 0.114. Point H Regions of Interest**
**Regions 1, 4: parabasal layer**
**Region 2: basal layer**
**Region 3: basement membrane**
**Regions 5, 6, 7: columnar epithelium**
**Regions 8, 9: stratified squamous epithelium, epithelial surface**

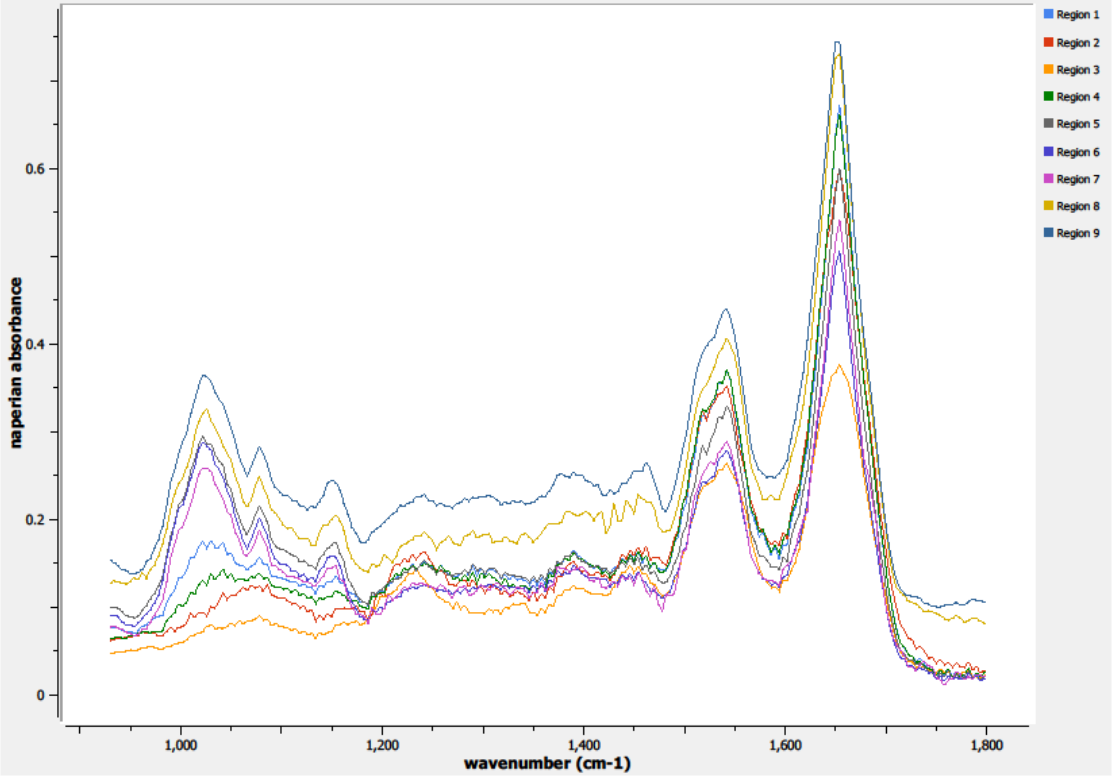



**Figure 0.115. Point H Spectra**

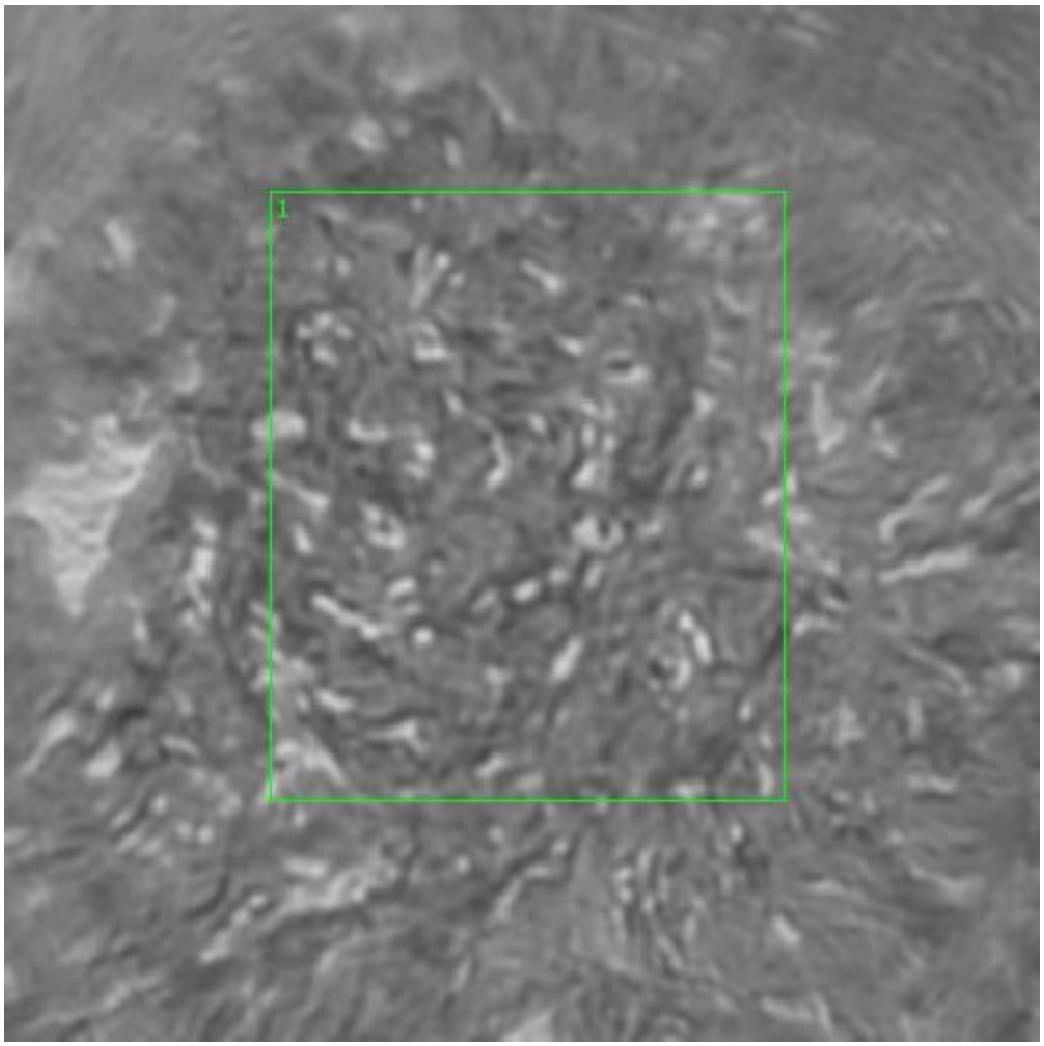



**Figure 0.116. Point I Regions of Interest**
**Region 1: basement membrane, lamina propia**

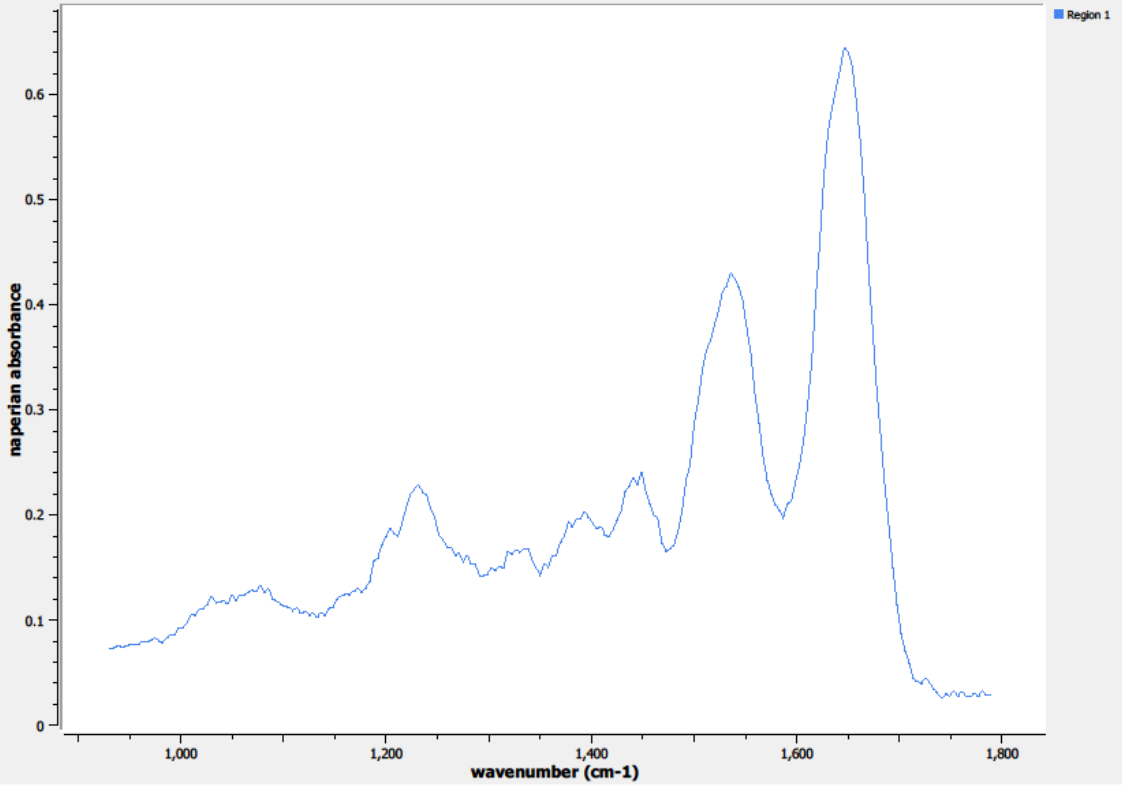



**Figure 0.117. Point I Spectra**

## A16. 12-05-A146a

| | |
|---|---|
| Unique ID | 12-05-A146a |
| Age | 69 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | none |

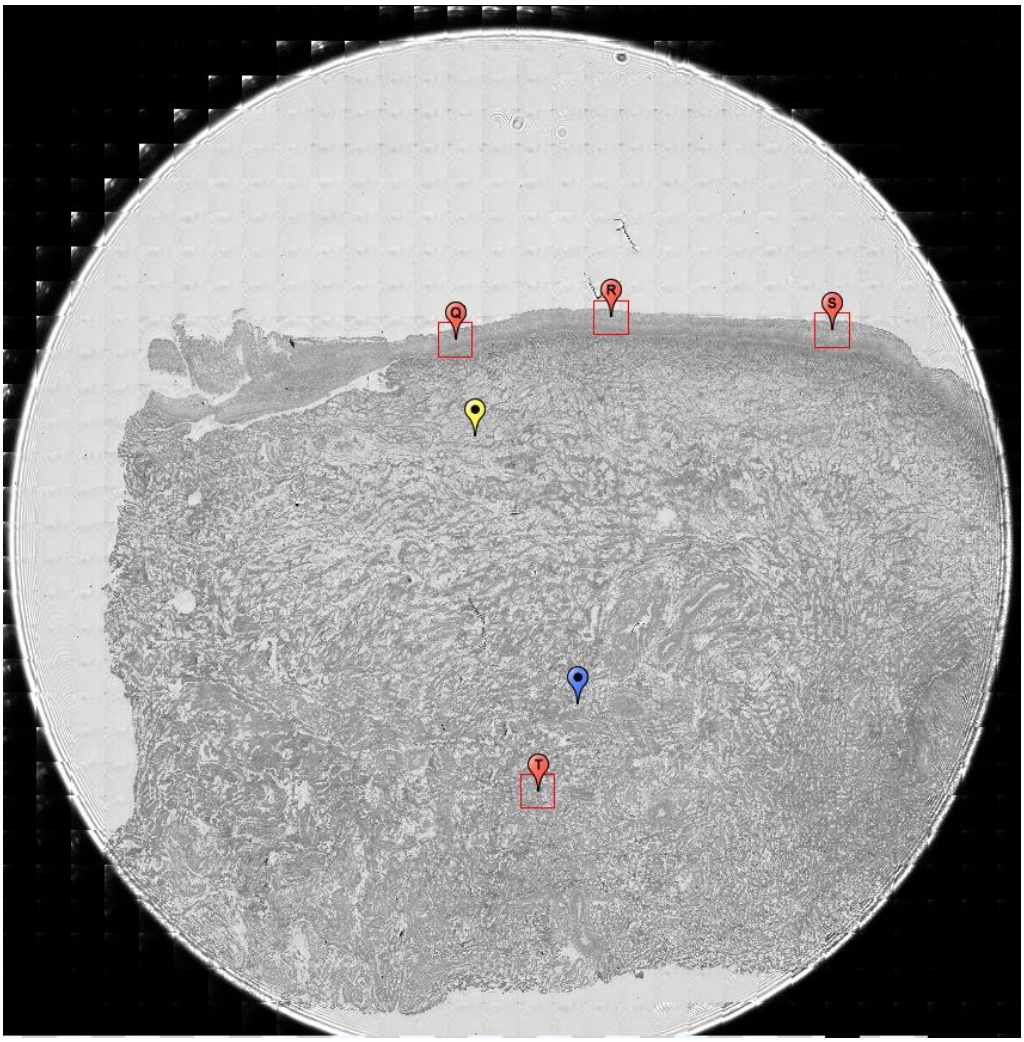



**Figure 0.118. Specimen 12-05-A146a, normal cervical tissue.**

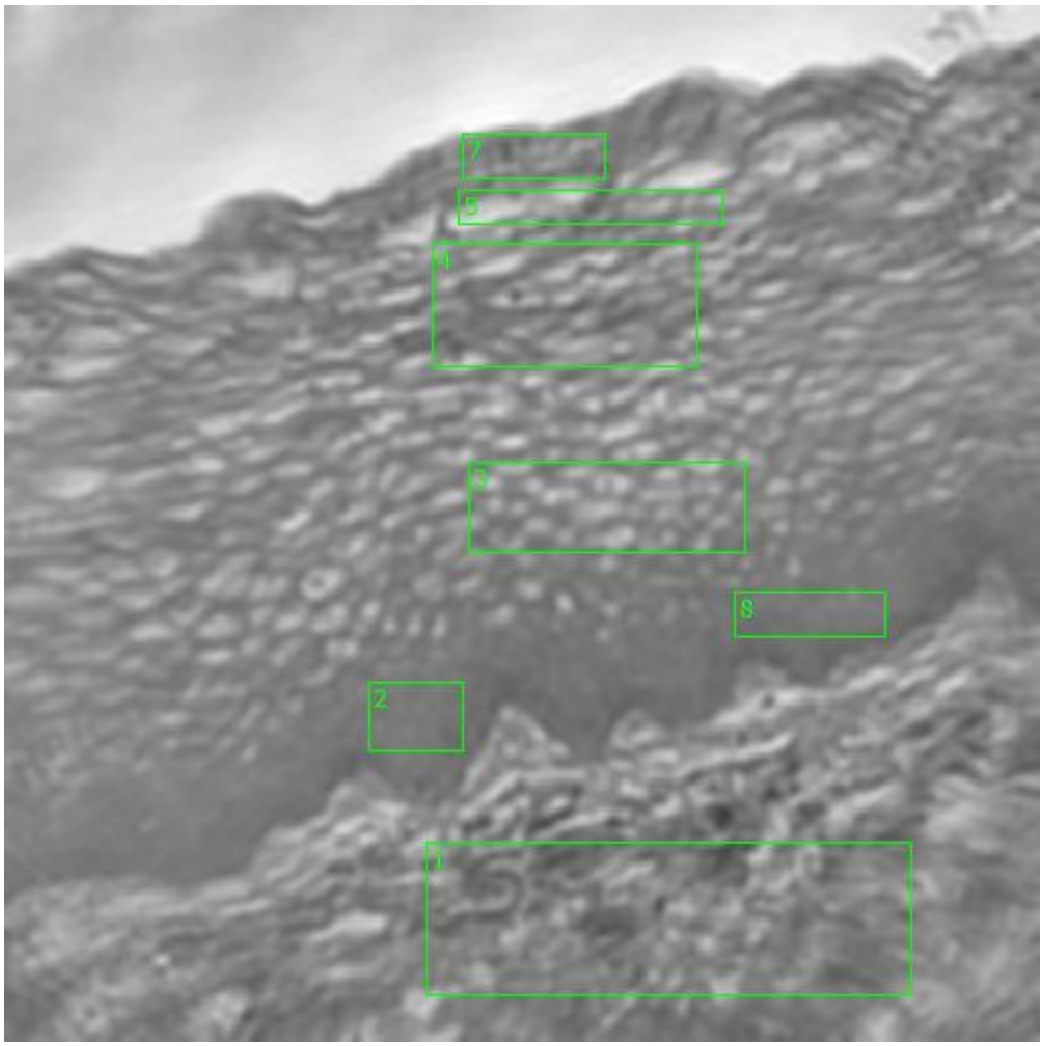



**Figure 0.119. Point Q Regions of Interest**
**Region 1: basement membrane**
**Regions 2, 8: basal layer**
**Region 3: columnar epithelium, parabasal layer**
**Regions 4, 5: stratified squamous epithelium**
**Region 7: epithelial surface**

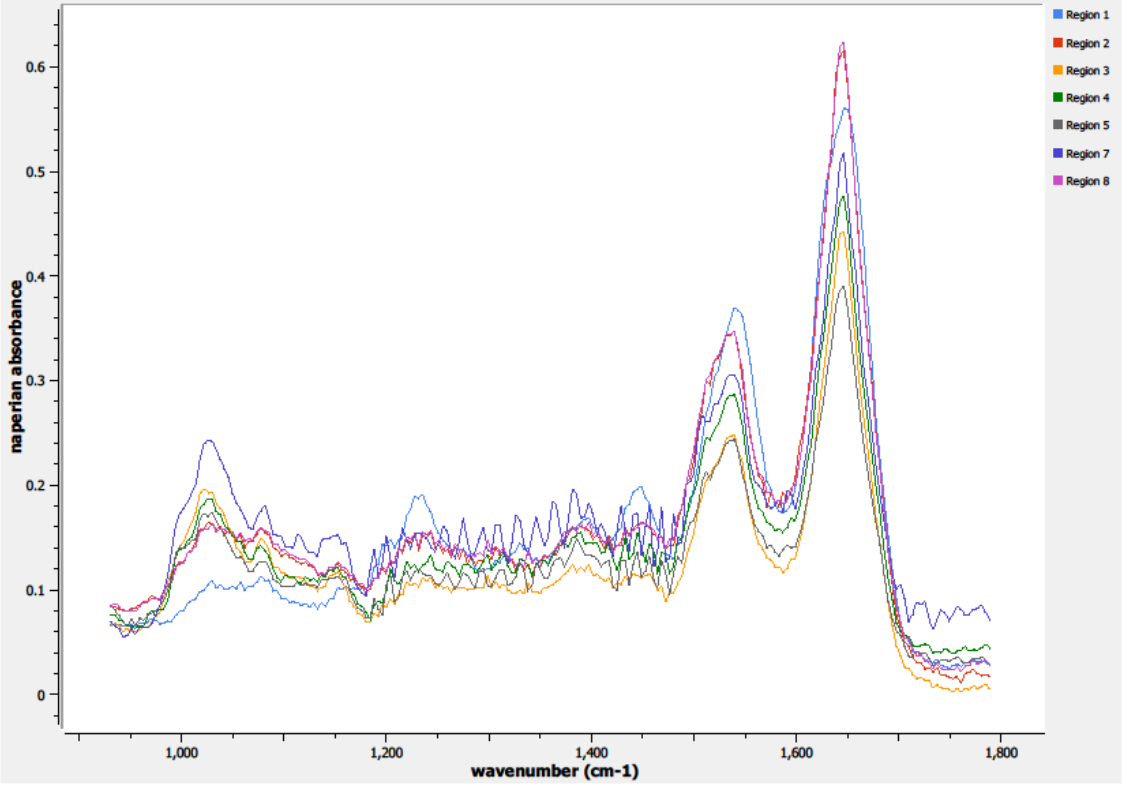



**Figure 0.120. Point Q Spectra**

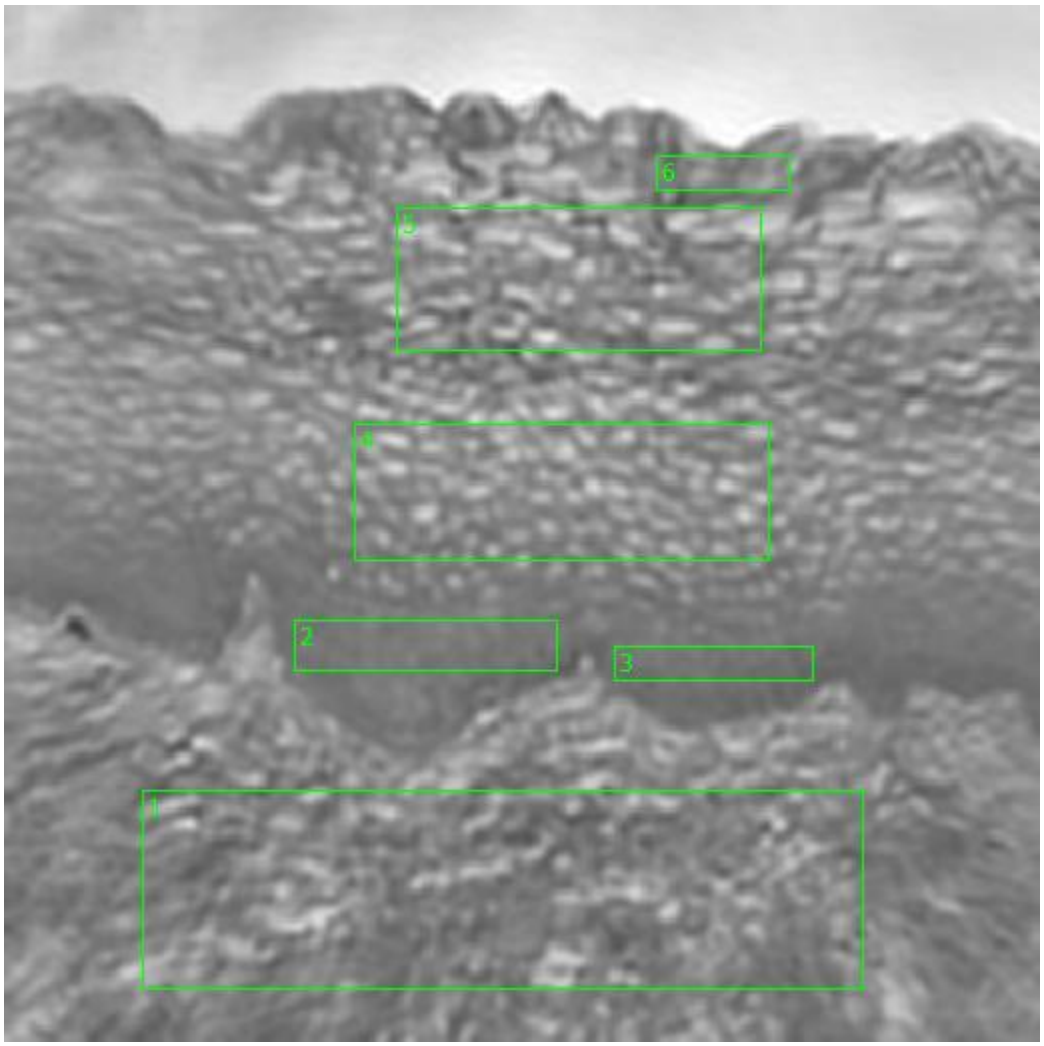



**Figure 0.121. Point R Regions of Interest**

**Region 1: basement membrane**
**Regions 2, 3: basal layer**
**Region 4: columnar epithelium, parabasal layer**
**Region 5: stratified squamous epithelium, intermediate**
**Region 6: epitherlial surface**

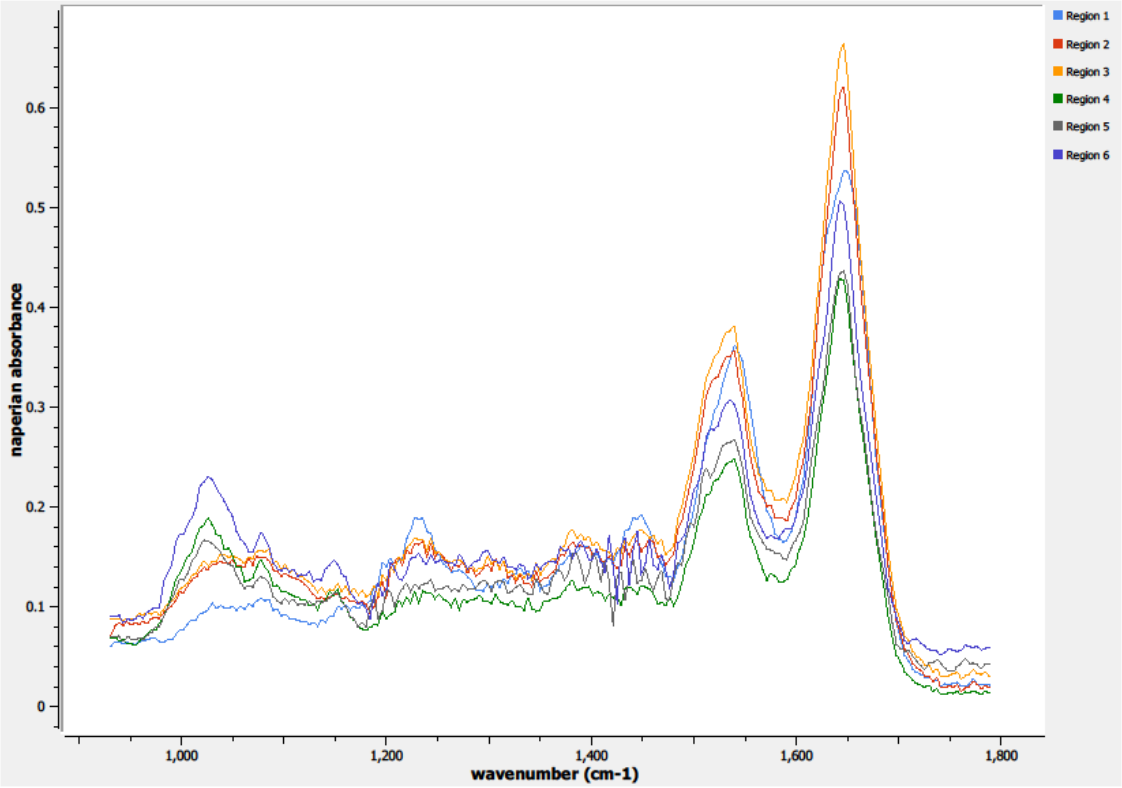



**Figure 0.122. Point R Spectra**

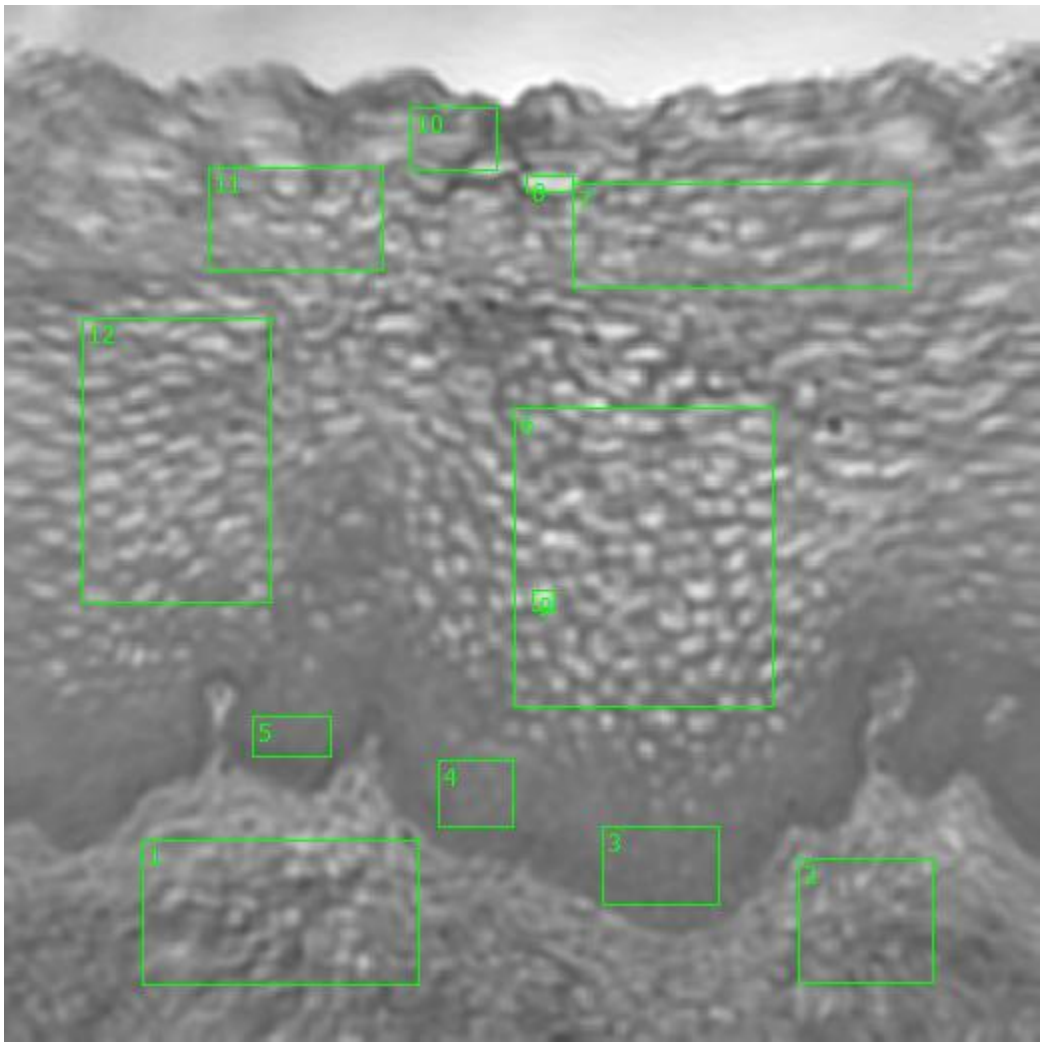



**Figure 0.123. Point S Regions of Interest**
**Regions 1, 2: basement membrane**
**Regions 3, 4, 5: basal layer**
**Regions 6, 12: columnar epithelium, parabasal layer**
**Region 9: individual columnar epithelial cell**
**Regions 7, 10, 11: stratified squamous epithelium, intermediate**
**Region 8: individual stratified squamous epithelial cell**

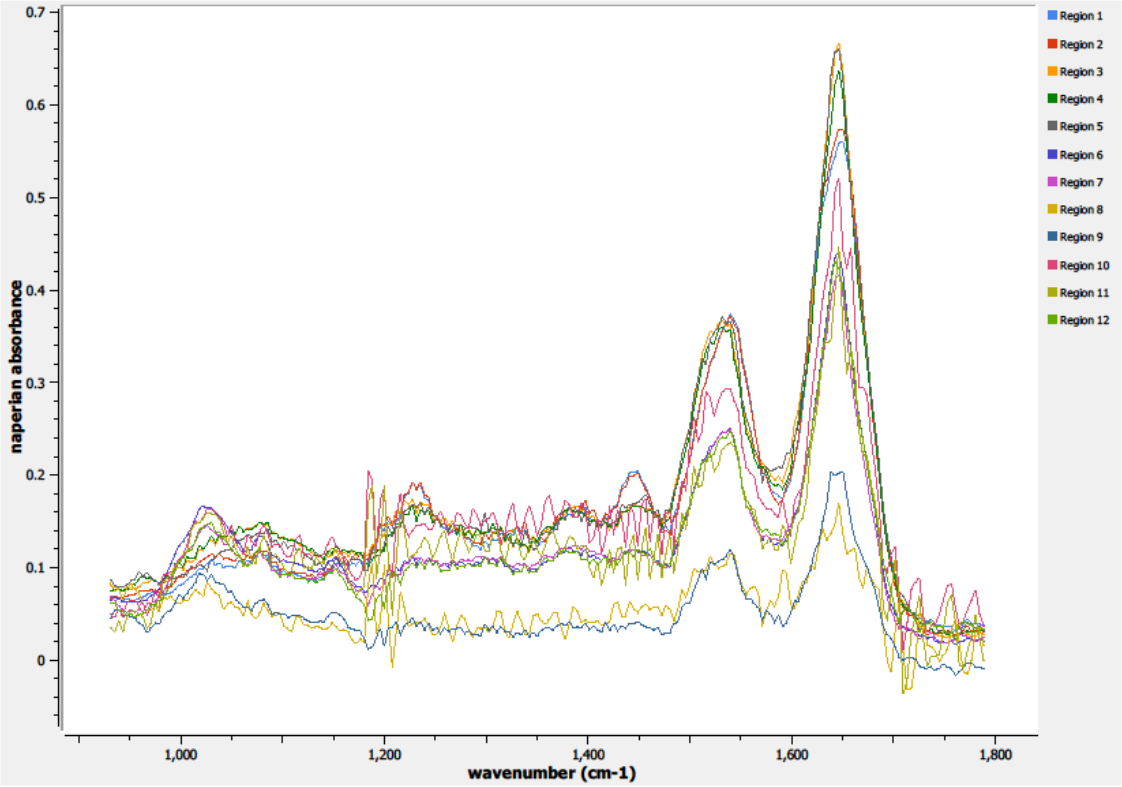



**Figure 0.124. Point S Spectra**

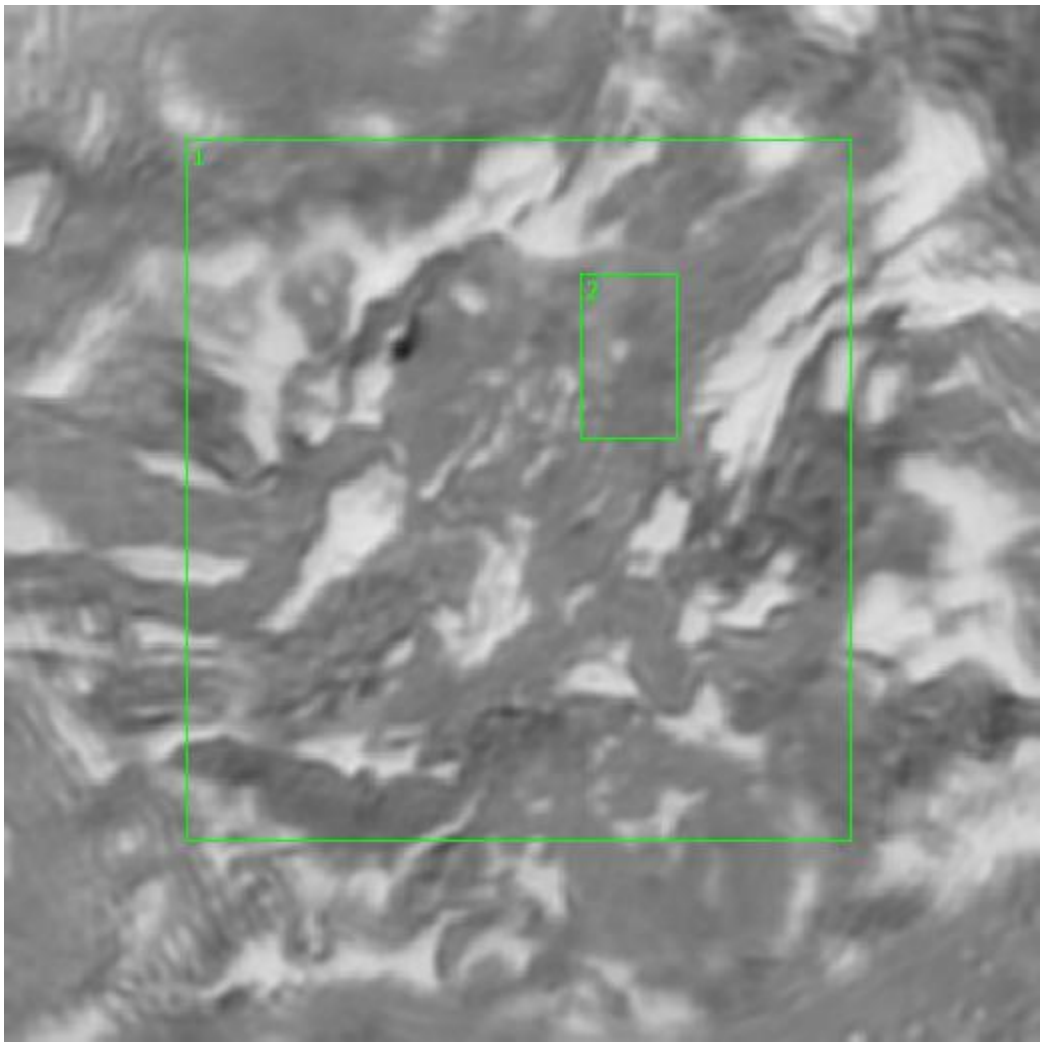



**Figure 0.125. Point T Regions of Interest**
**Regions 1, 2: lamina propia**

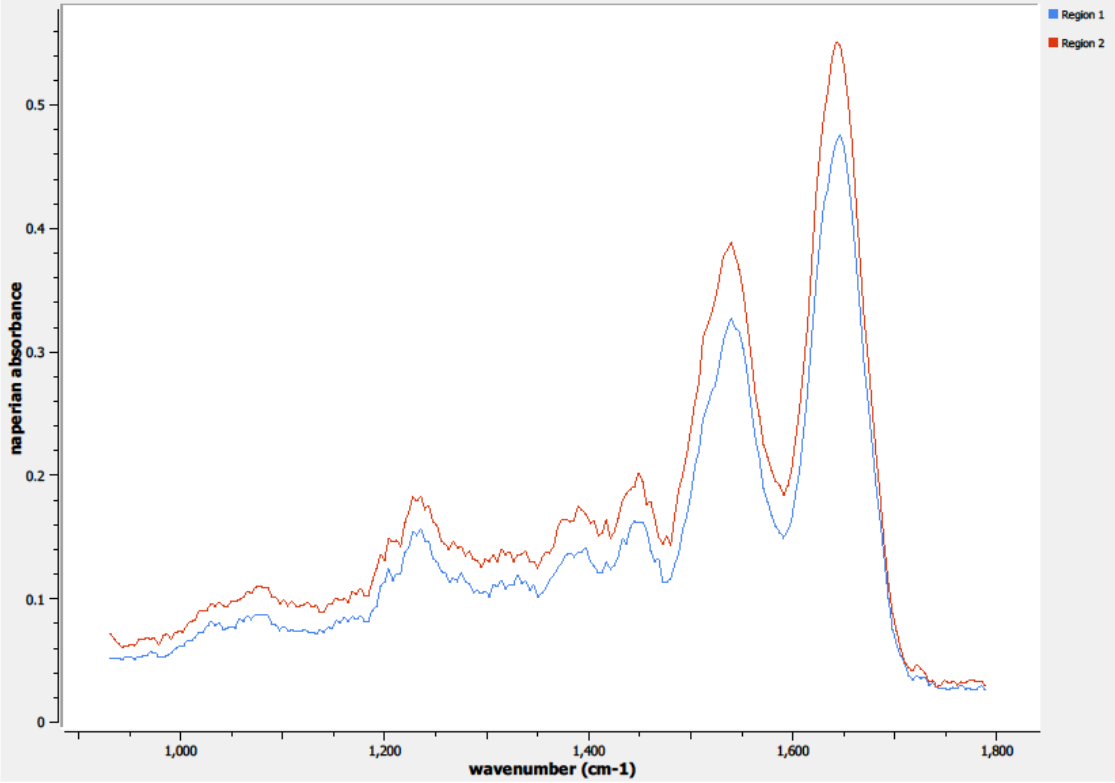



**Figure 0.126. Point T Spectra**

# Appendix B

# Cervical Cancer

**A1.01-09-A245a**

| | |
|---|---|
| Unique ID | 01-09-A245a |
| Age | 64 |
| Race | black |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix tumor, invasive squamous cell carcinoma, %T (25) / %N (10), may not use |

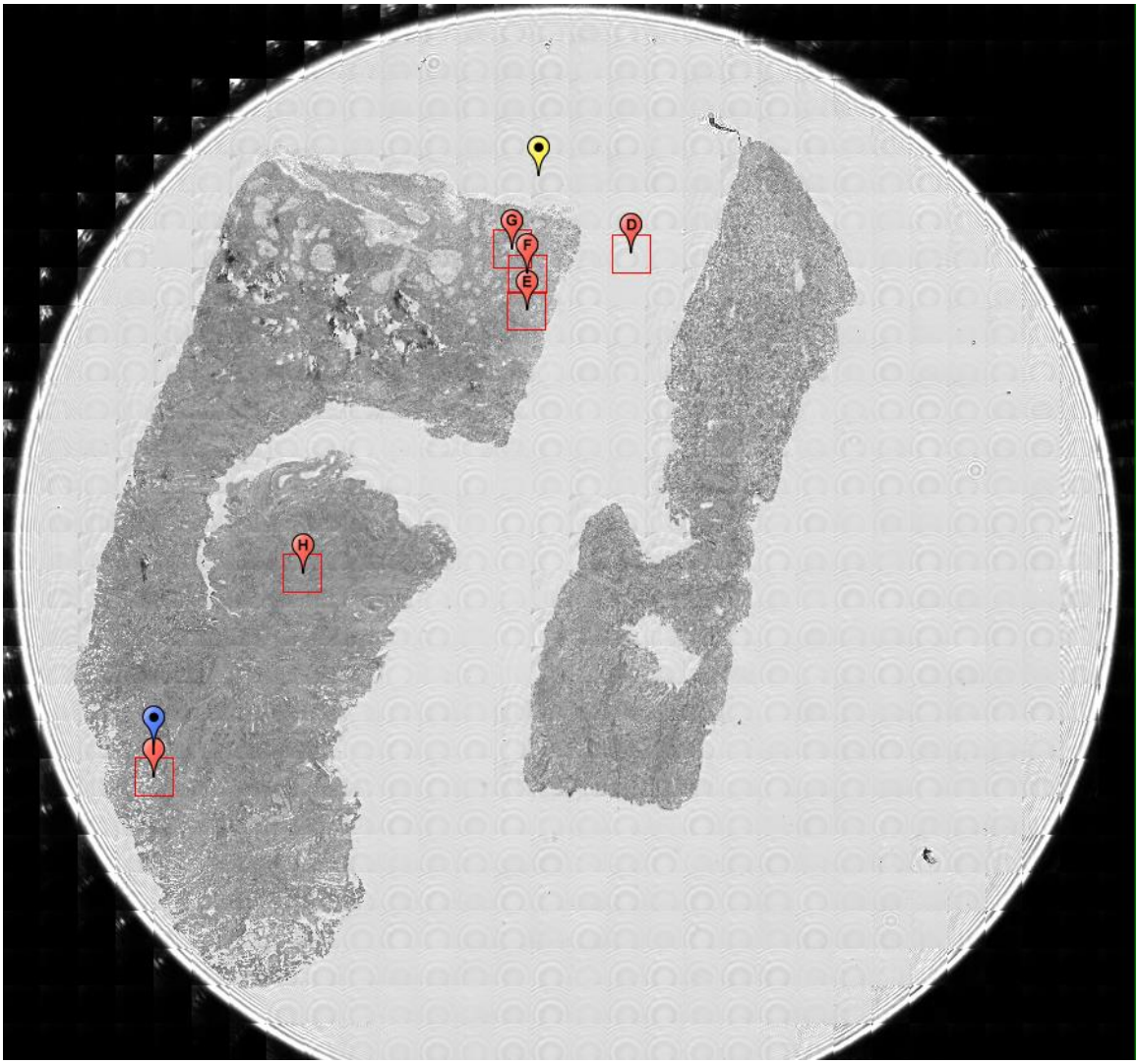



**Figure 0.1. Specimen 01-09-A245a, cervical cancer tissue**

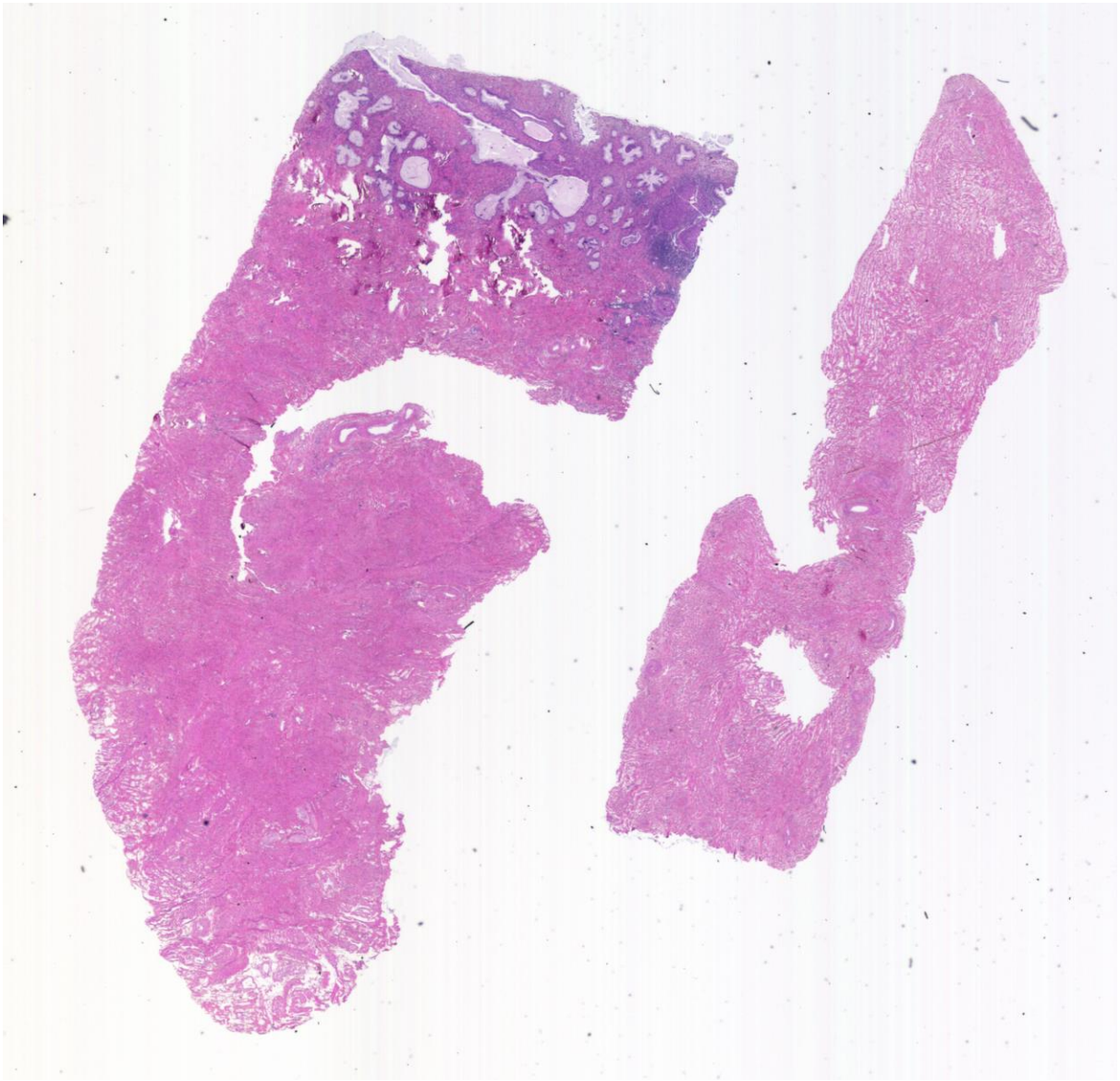


**Figure 0.2. Specimen 01-09-A245a, H&E stain, squamous cell carcinoma of the cervix**

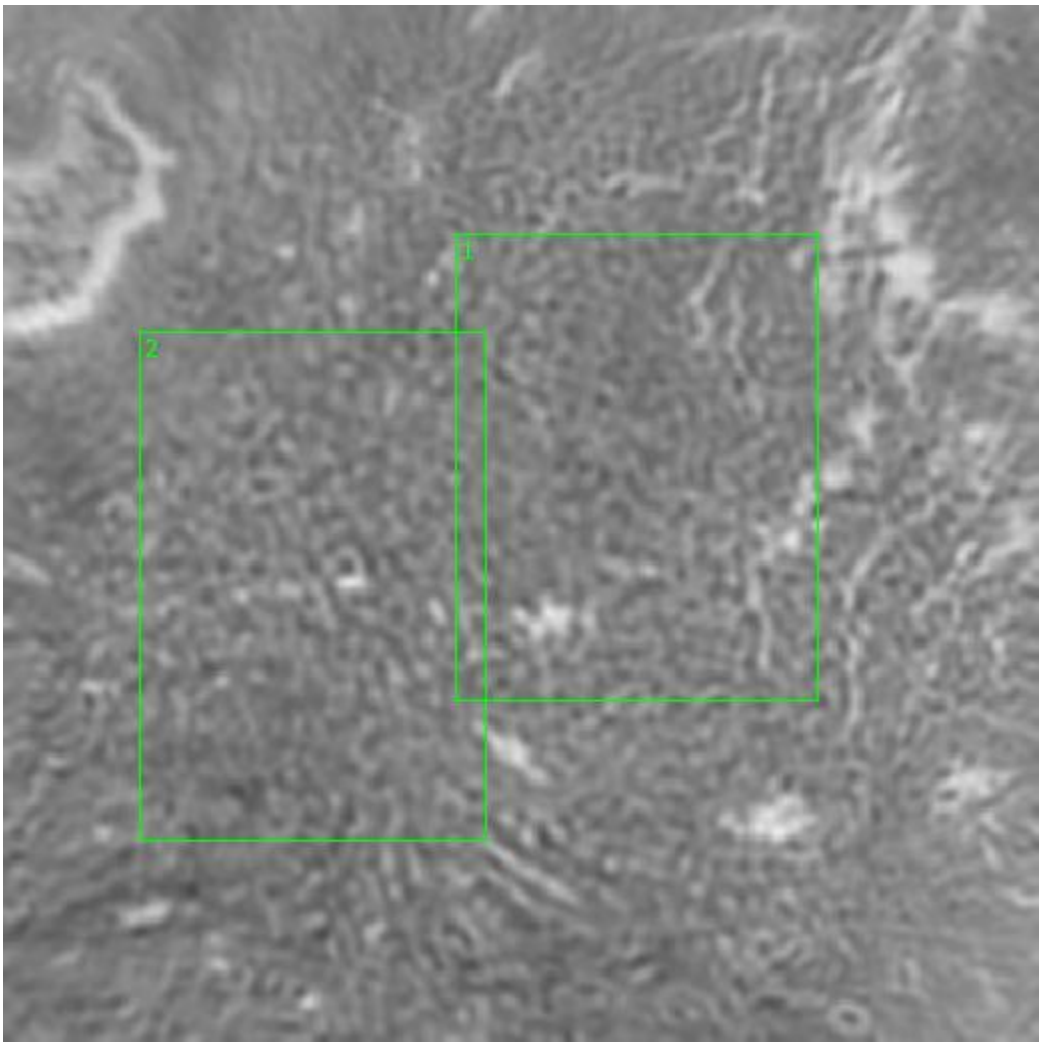



**Figure 0.3. Point E Regions of Interest**
**Regions 1, 2: lamina propia, squamous cell carcinoma**

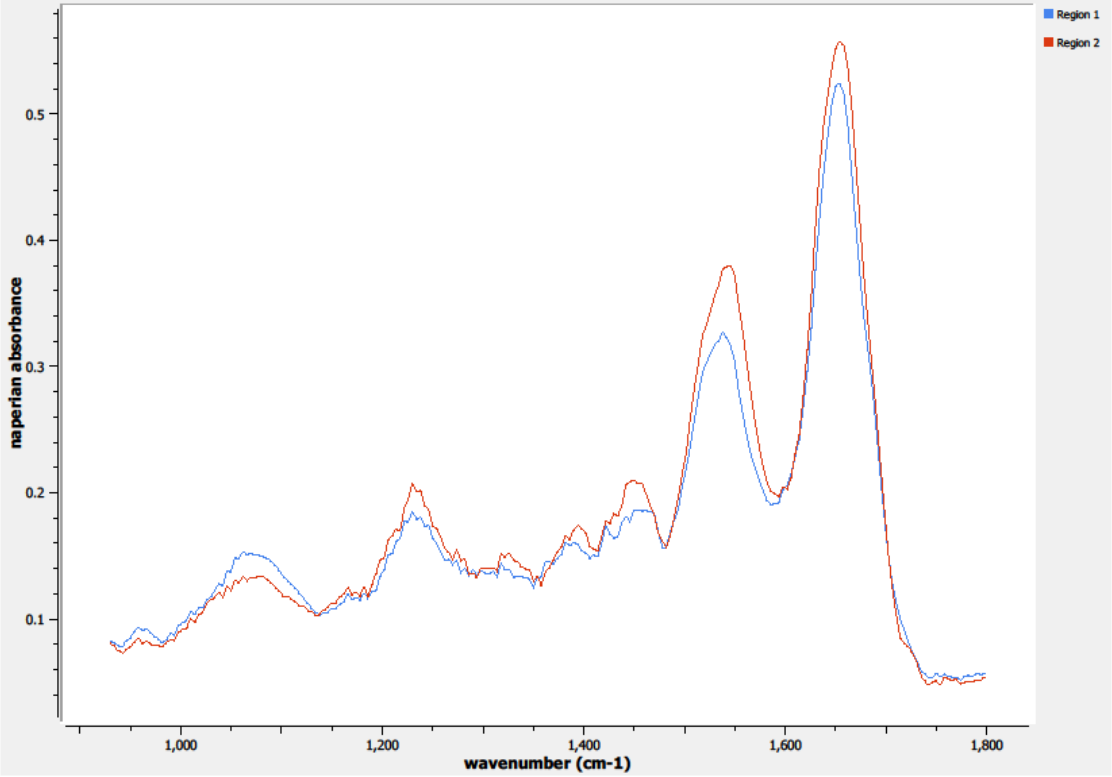



**Figure 0.4. Point E Spectra**

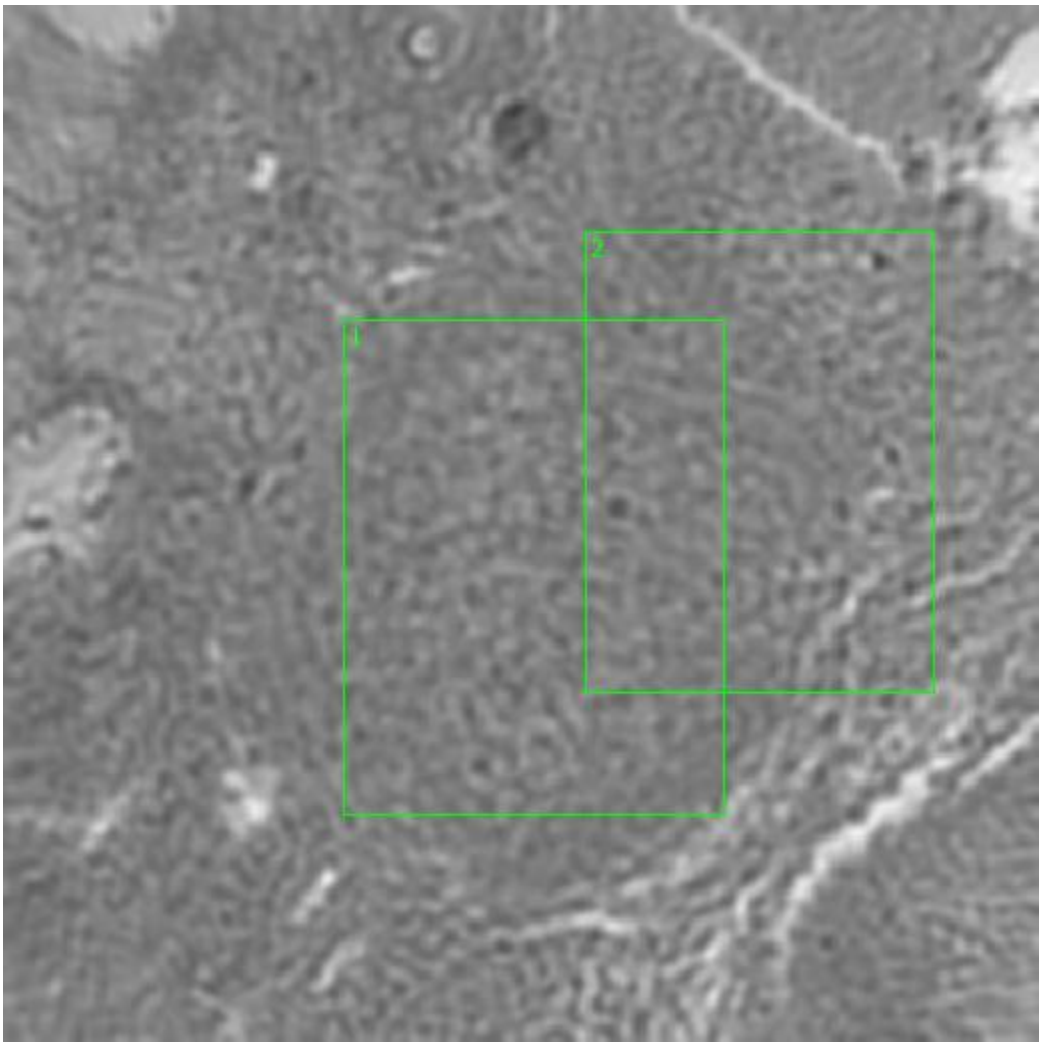



**Figure 0.5. Point F Regions of Interest**
**Regions 1, 2: lamina propia, squamous cell carcinoma**

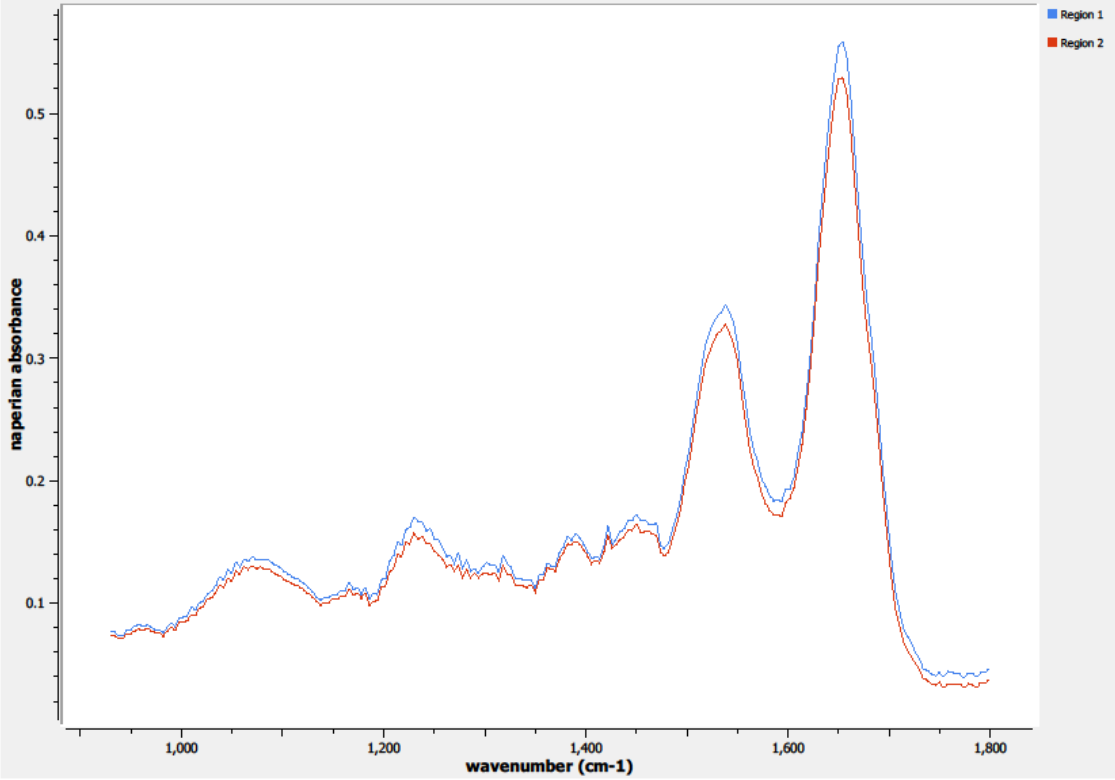



**Figure 0.6. Point F Spectra**

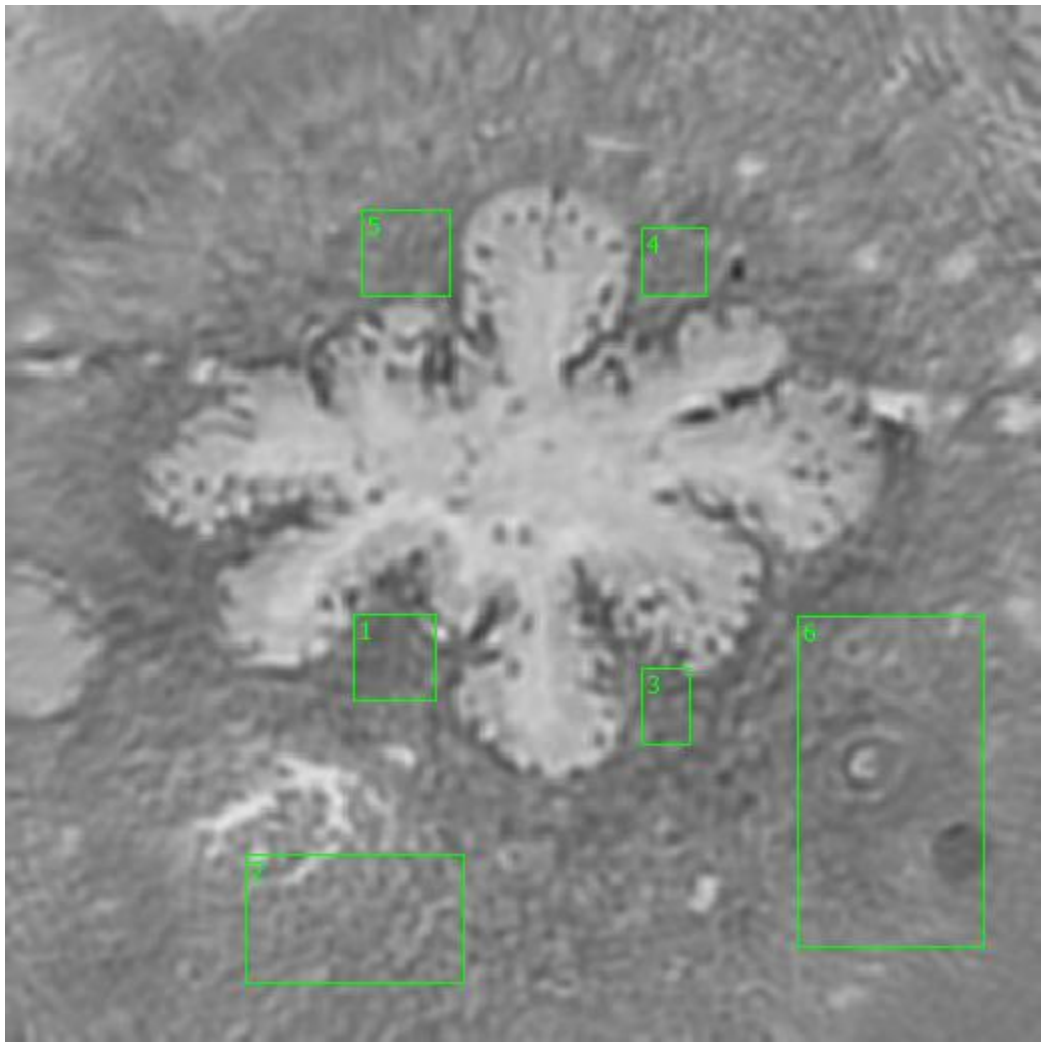



**Figure 0.7. Point G Regions of Interest**

**Regions 1, 3, 4, 5: edge of glandular structure, squamous cell carcinoma**

**Regions 2, 6: lamina propia near glandular structure, squamous cell carcinoma**

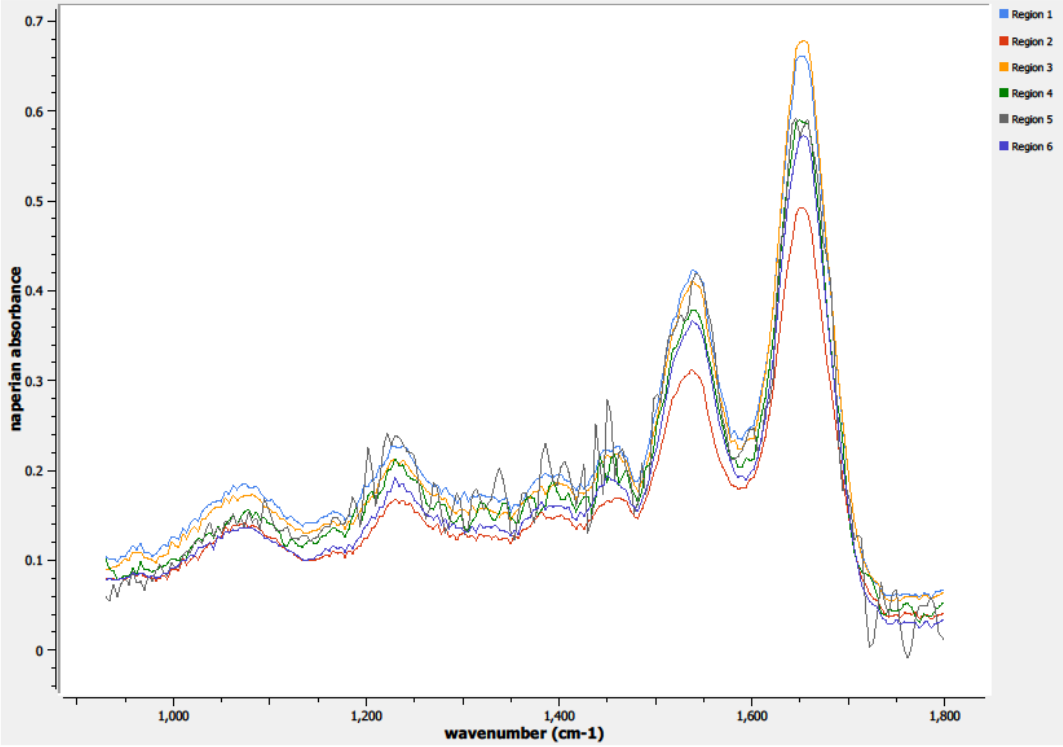



**Figure 0.8. Point G Spectra**

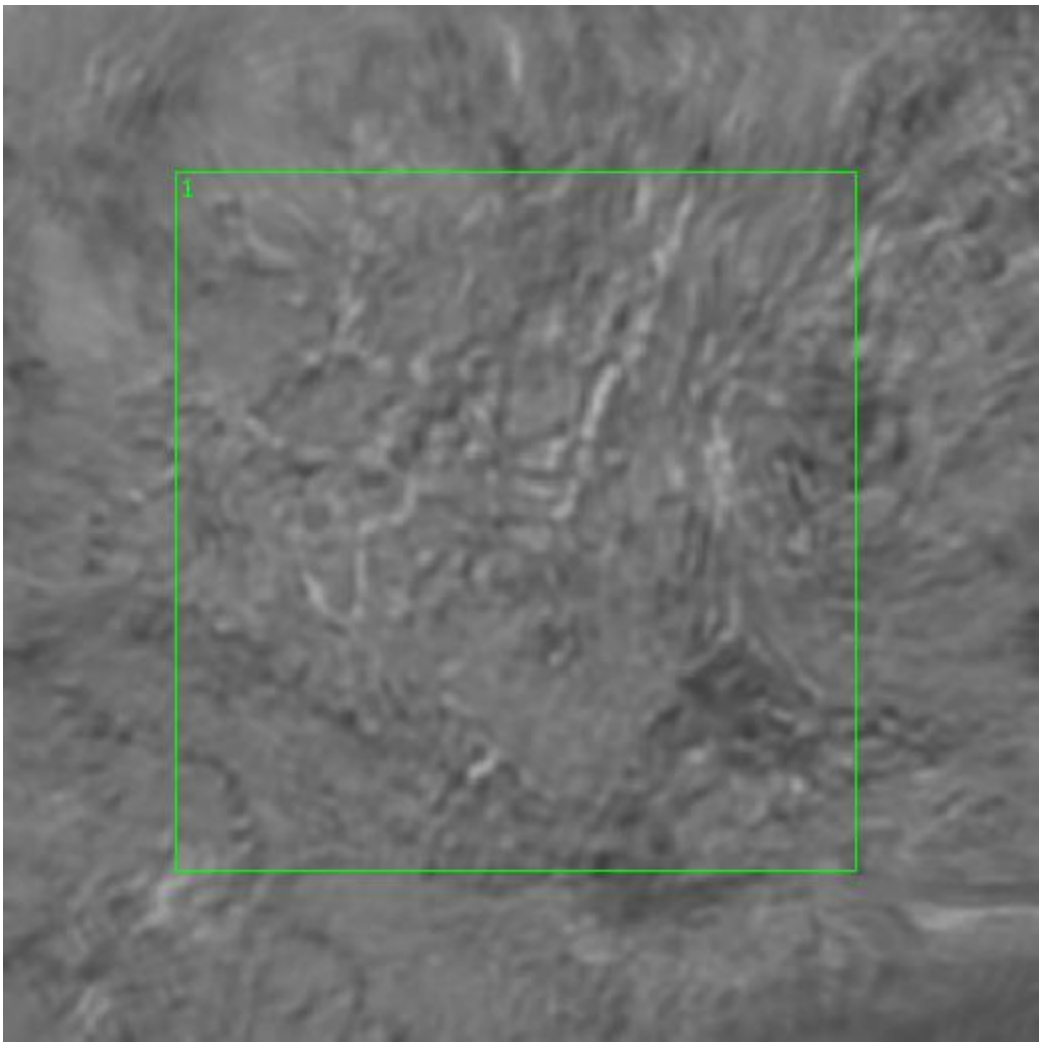



**Figure 0.9. Point H Regions of Interest**
**Region 1: lamina propia, transition region between squamous cell carcinoma and normal tissue**

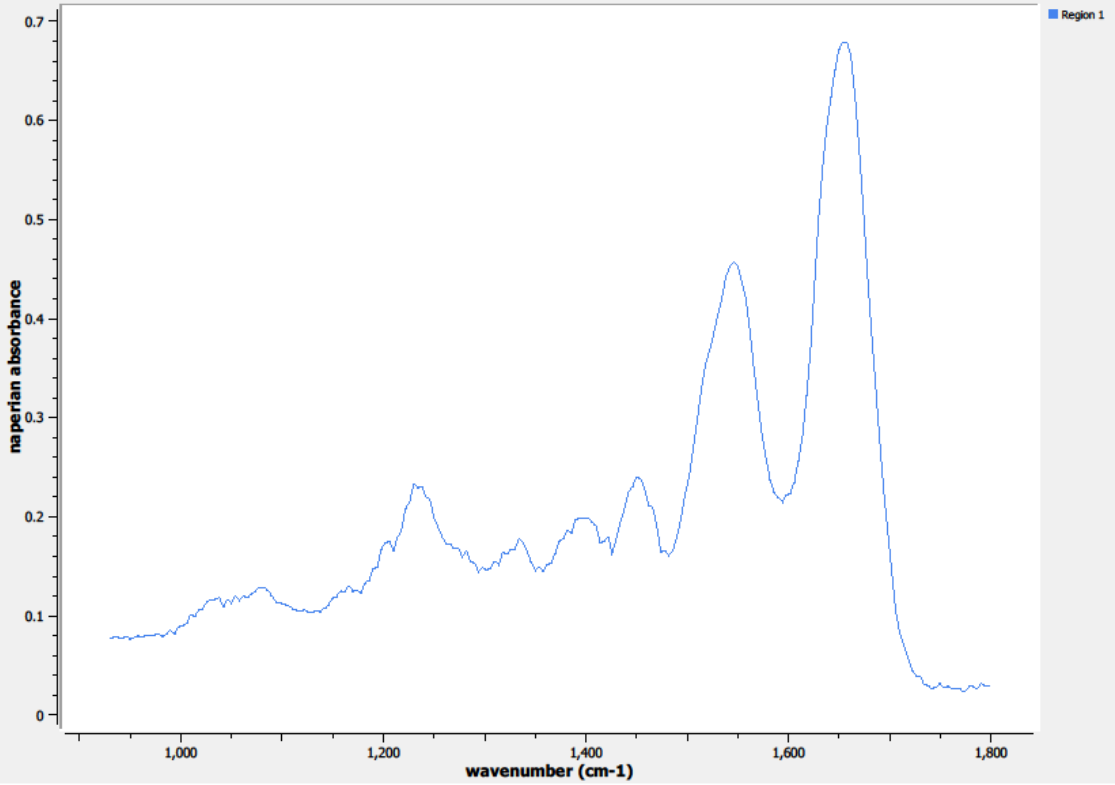



**Figure 0.10. Point H Spectra**

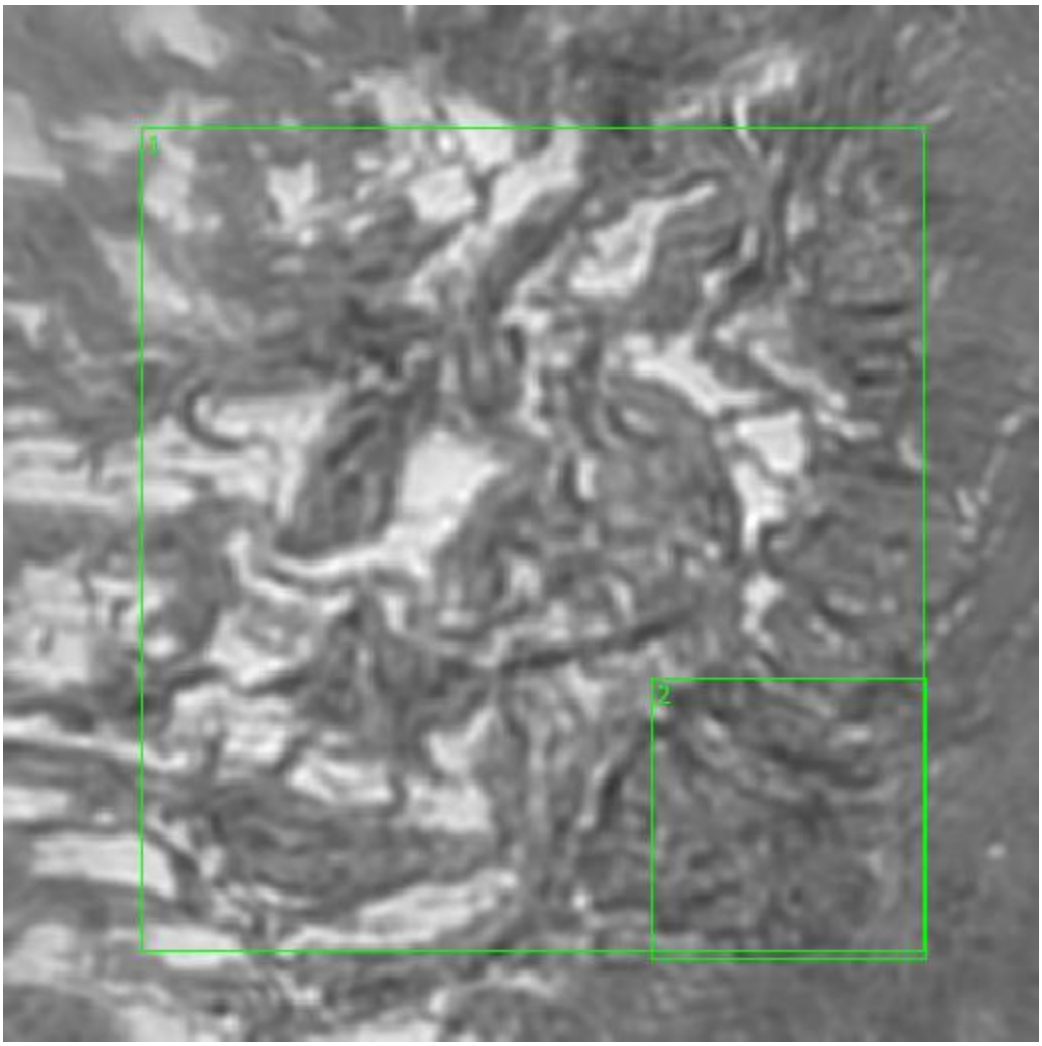



**Figure 0.11. Point I Regions of Interest**
**Regions 1, 2: lamina propia, transition region between squamous cell carcinoma and normal tissue**

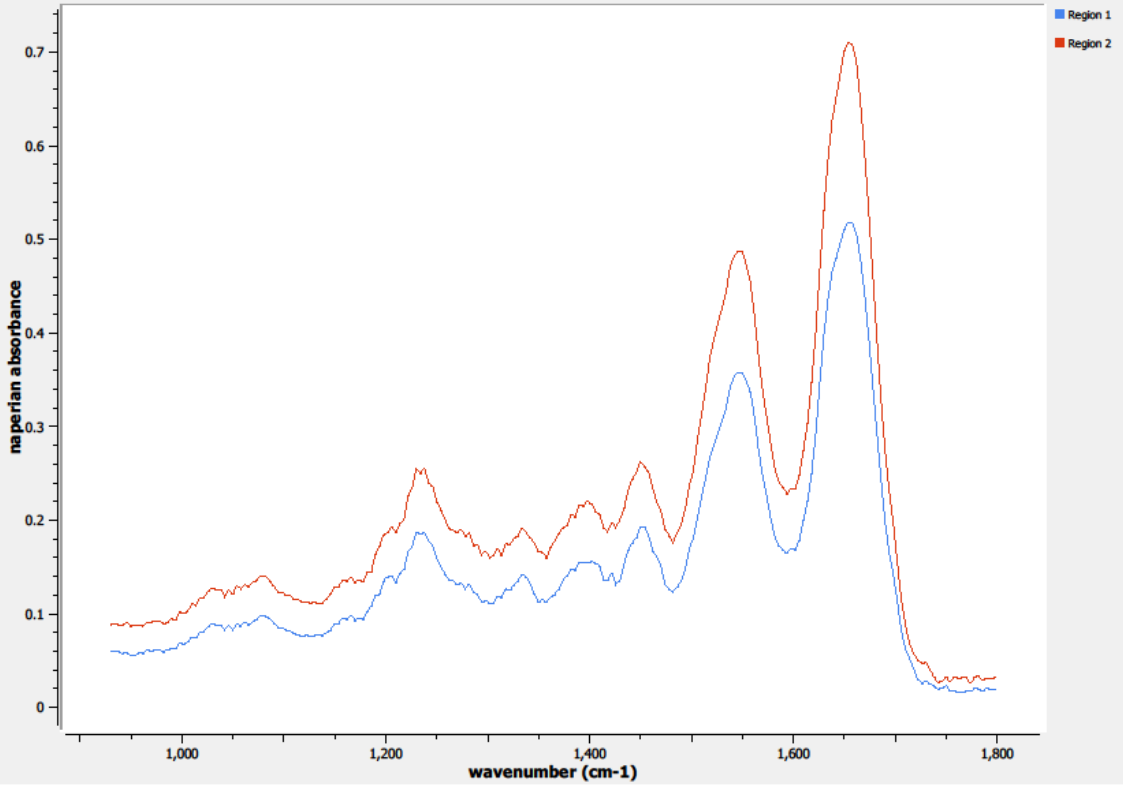



**Figure 0.12. Point I Spectra**

## A2.03-03-A268a

| | |
|---|---|
| Unique ID | 03-03-A268a |
| Age | 38 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix tumor, not clear transition between normal and cancer tissue, not use data in results |

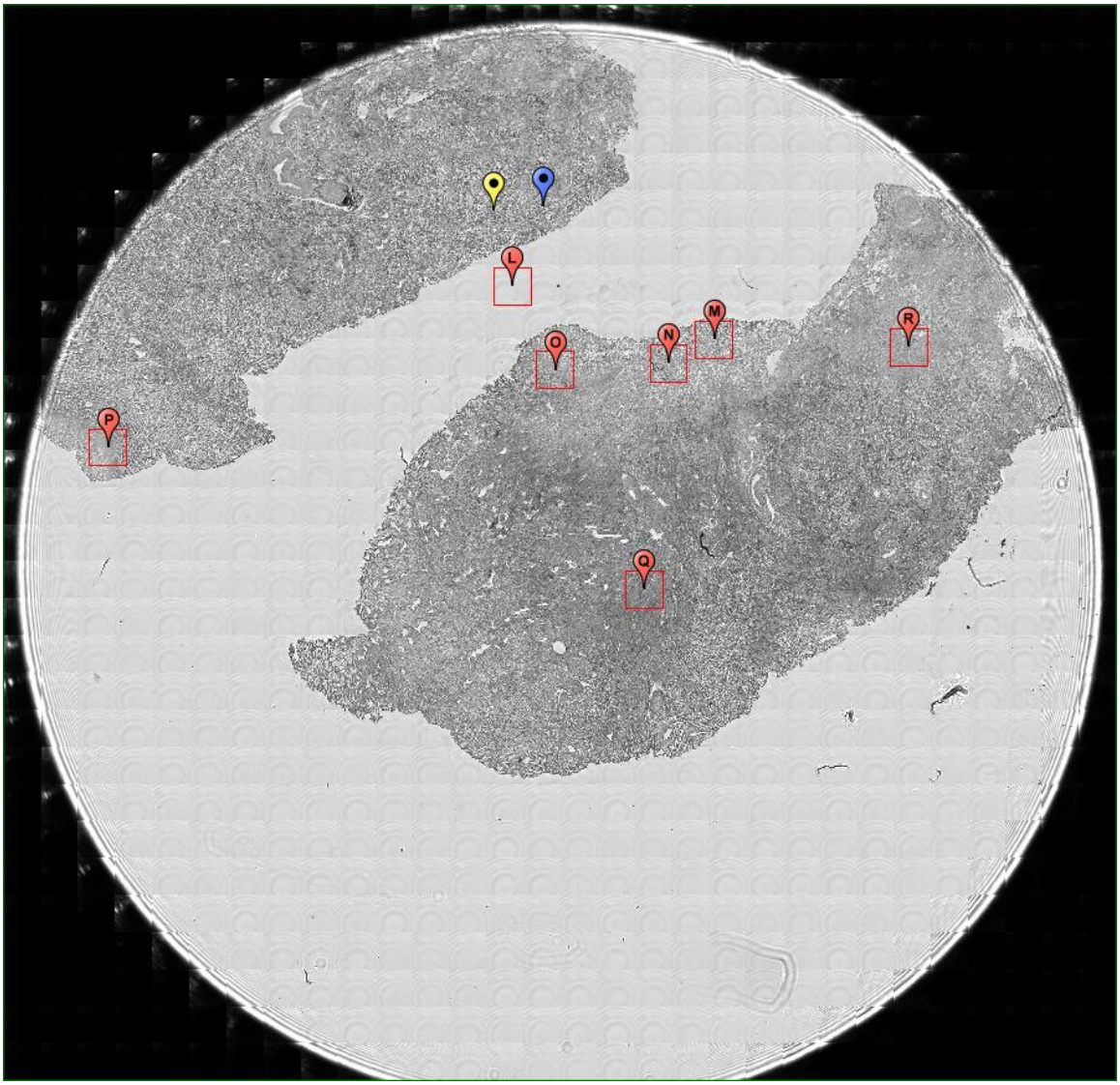



**Figure 0.13. Specimen 03-03-A268a, cervical cancer tissue**

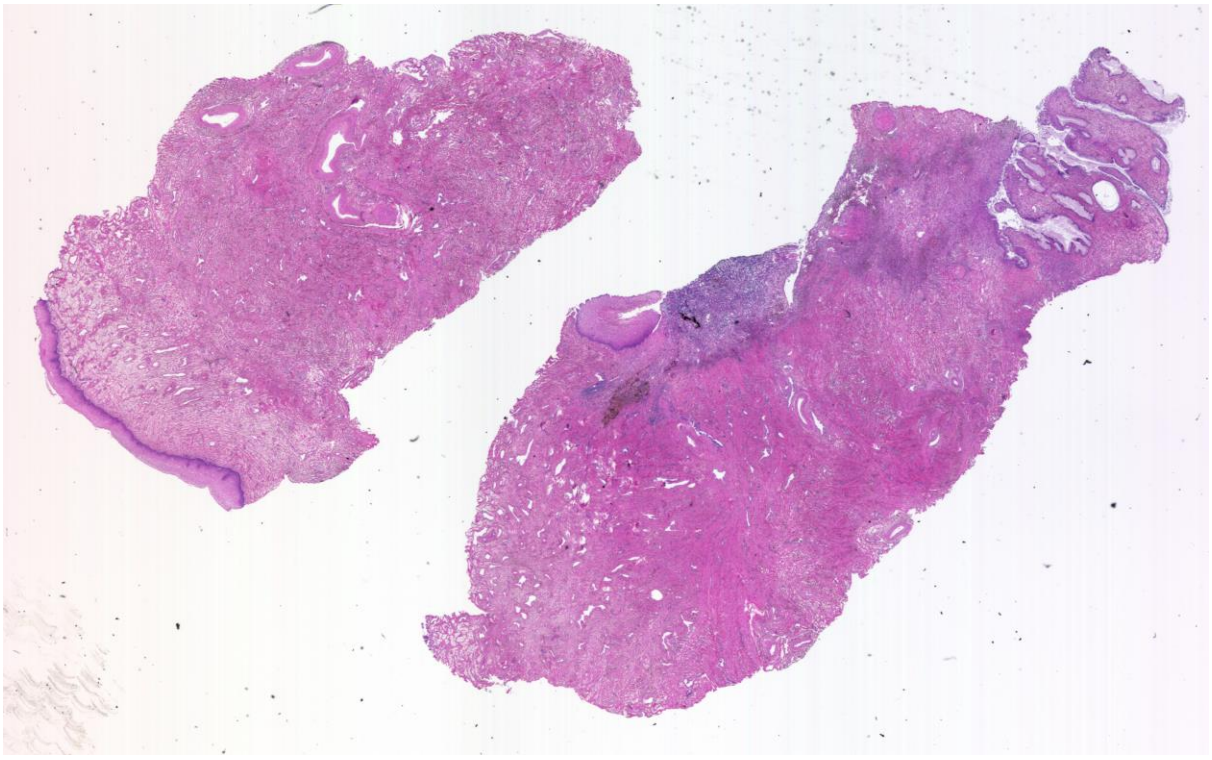


**Figure 0.14. Specimen 03-03-A268a, H&E stain, squamous cell carcinoma of the cervix**

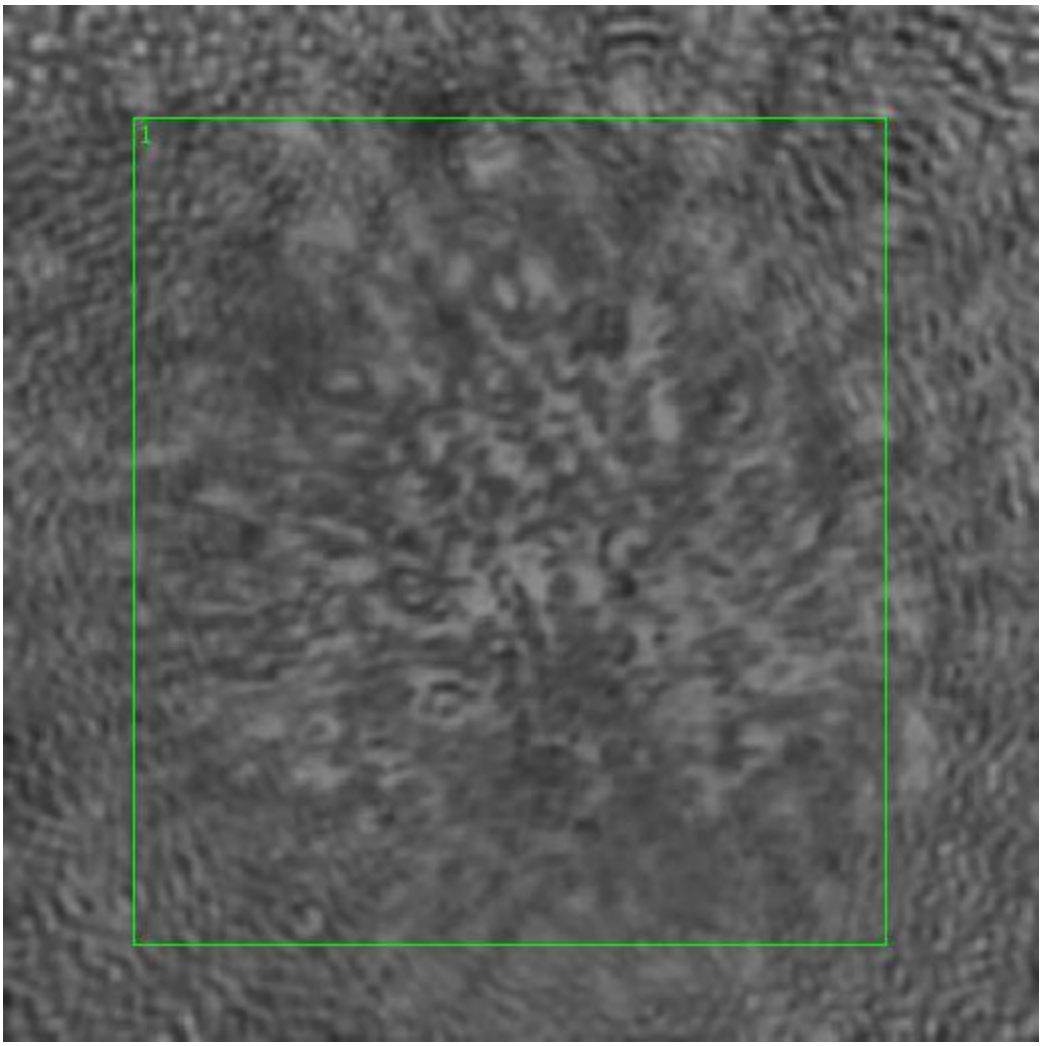



**Figure 0.15. Point M Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

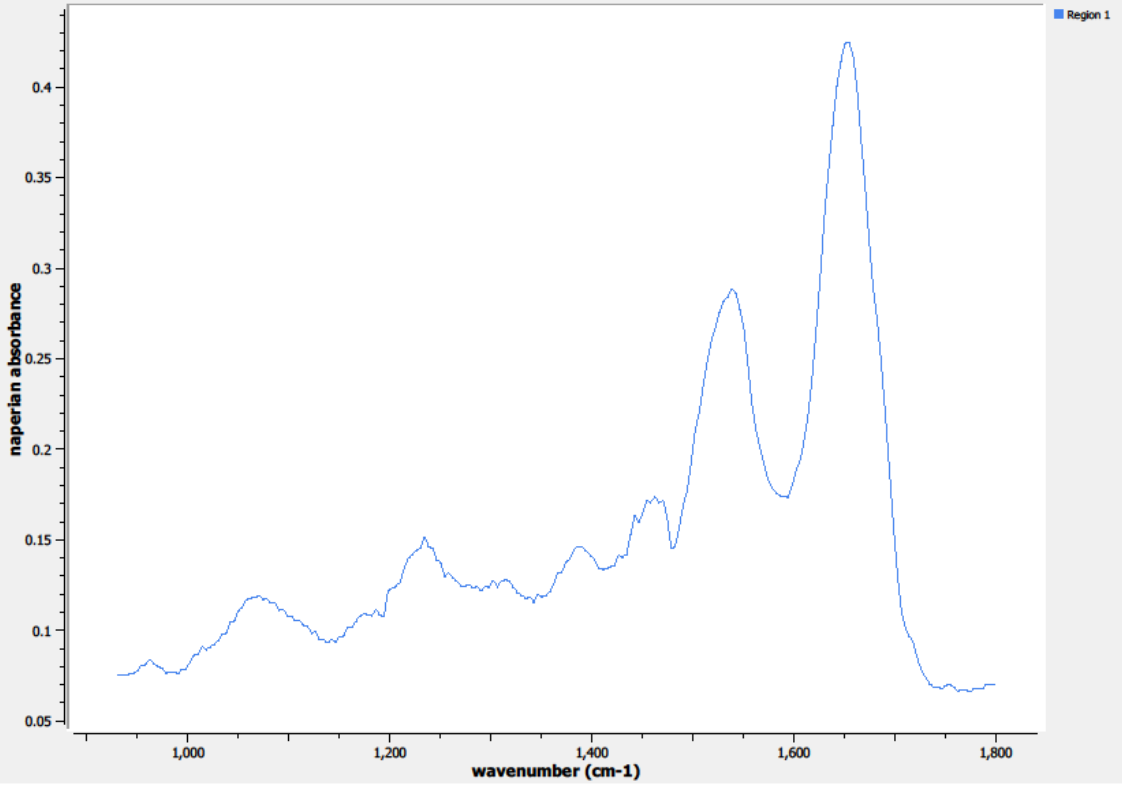



**Figure 0.16. Point M Spectra**

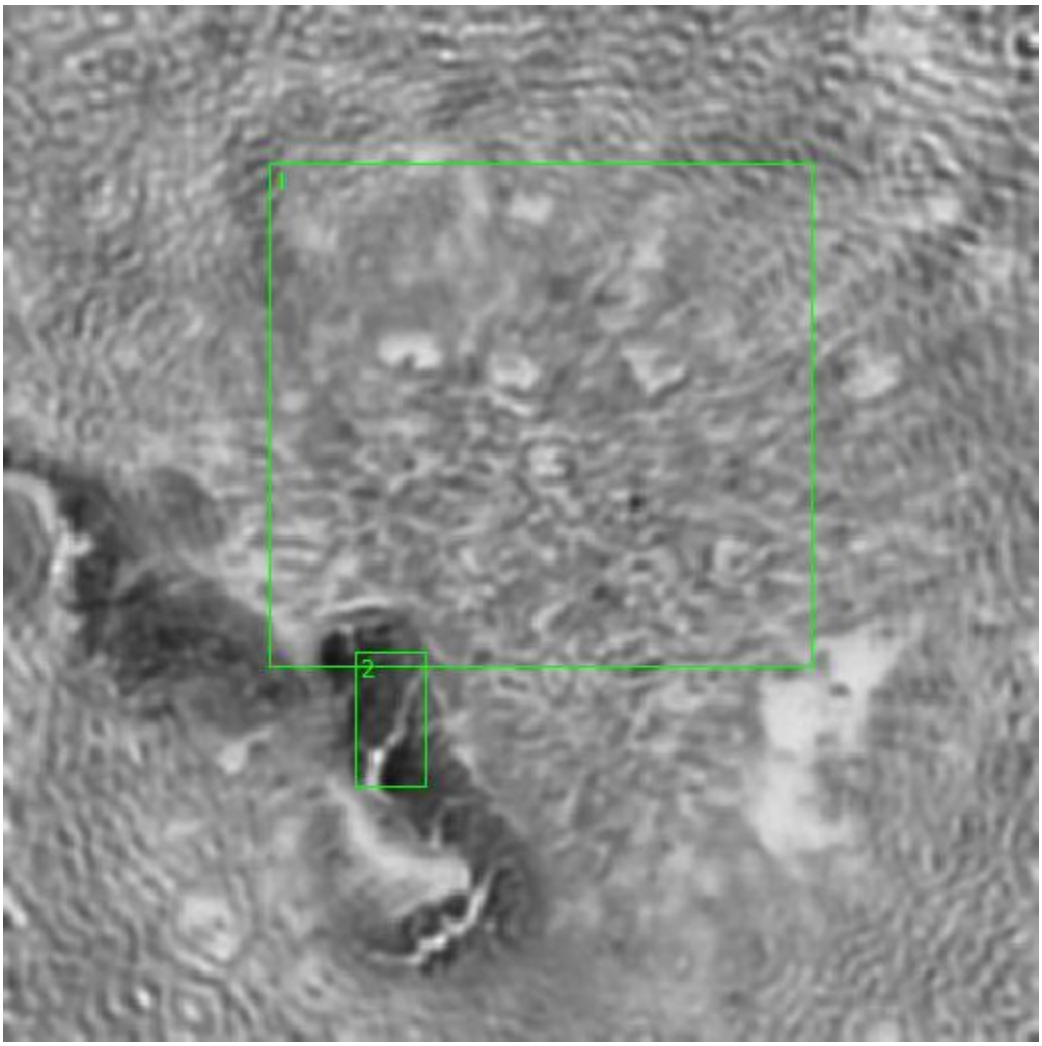



**Figure 0.17. Point N Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**
**Region 2: lamina propia, concentrated region, squamous cell carcinoma**

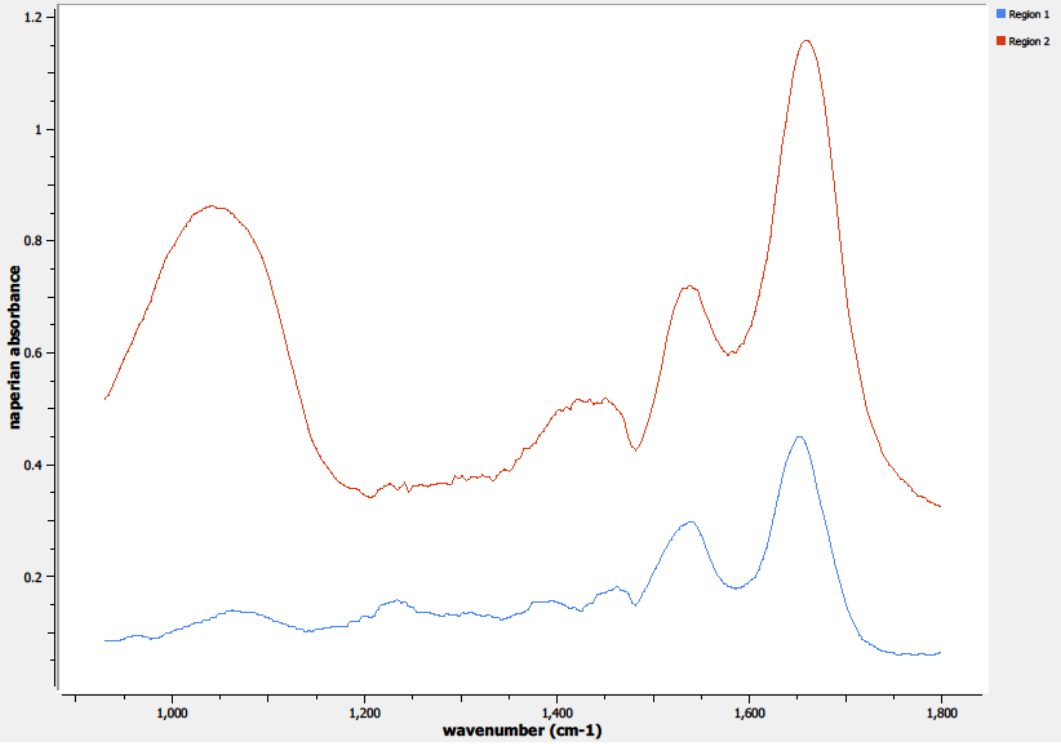



**Figure 0.18. Point N Spectra**

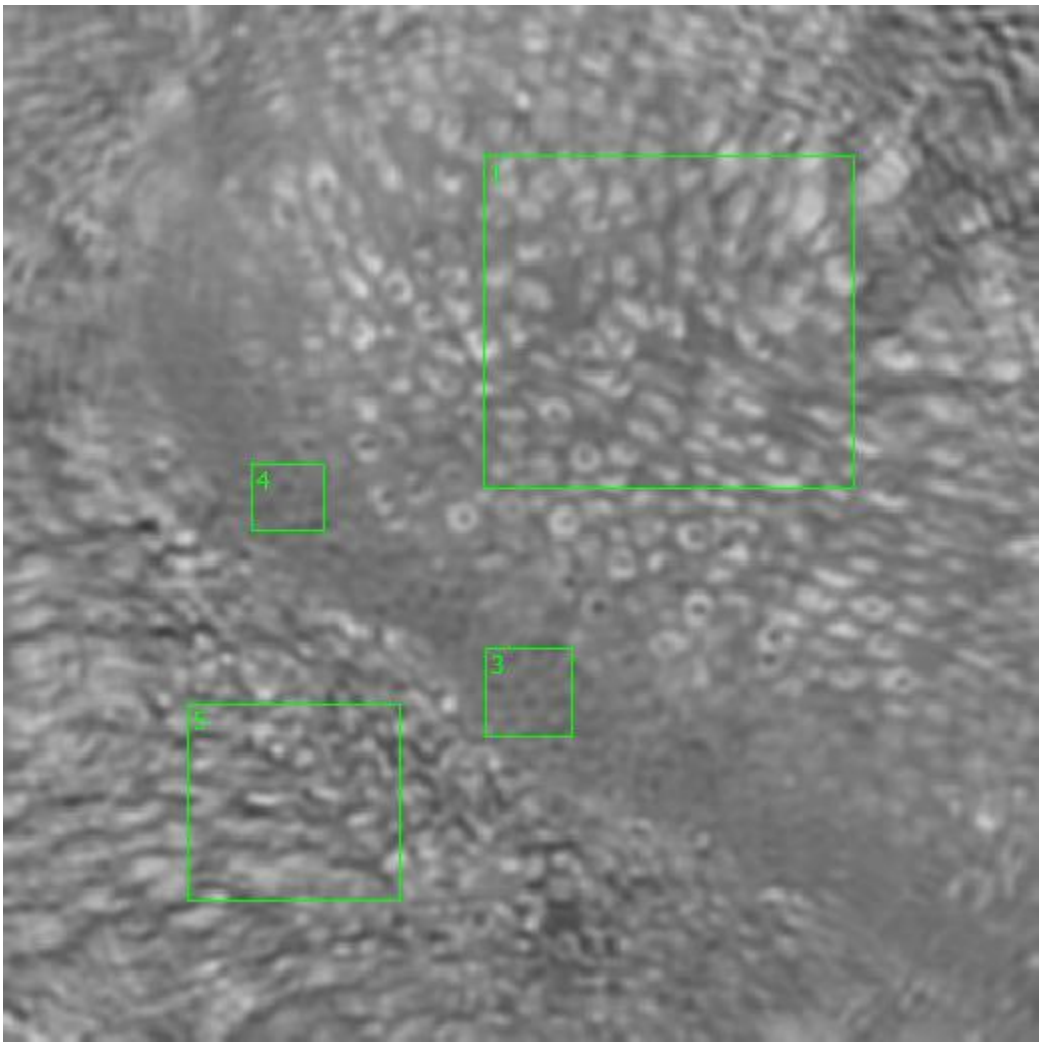



**Figure 0.19. Point O Regions of Interest**
**Region 1: columnar epithelium, transition region between squamous cell carcinoma and normal tissue**
**Regions 3, 4: basal layer, transition region between squamous cell carcinoma and normal tissue**
**Region 5: basement membrane, lamina propia, transition region between squamous cell carcinoma and normal tissue**

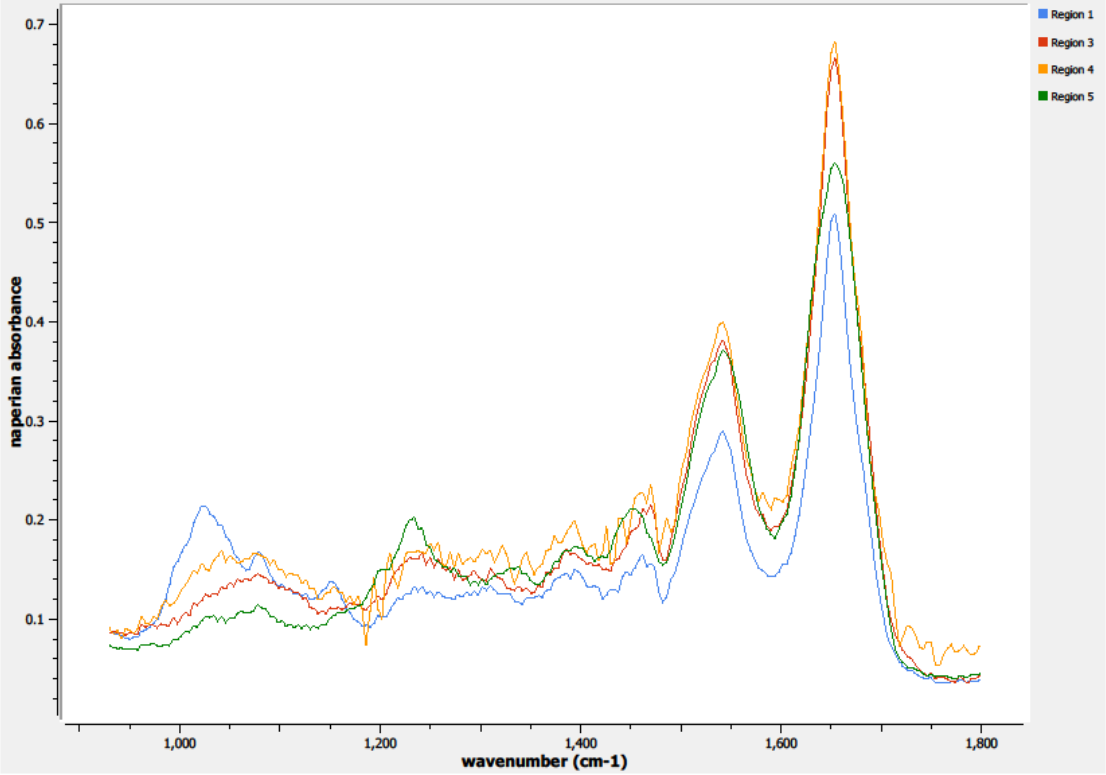



**Figure 0.20. Point O Spectra**

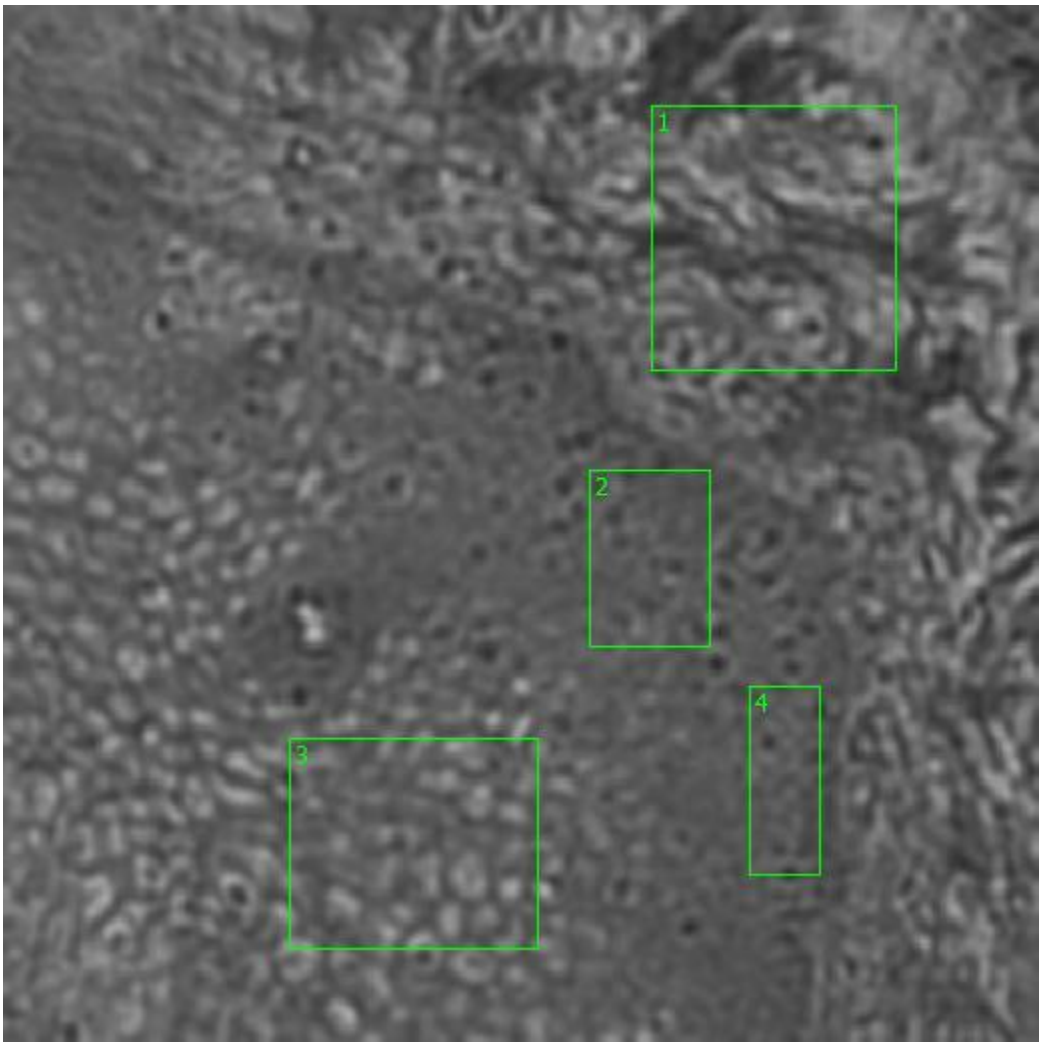



**Figure 0.21. Point P Regions of Interest**
**Region 1: basement membrane, lamina propia, transition region between squamous cell carcinoma and normal tissue**
**Regions 2, 4: basal layer, transition region between squamous cell carcinoma and normal tissue**
**Region 3: columnar epithelium, transition region between squamous cell carcinoma and normal tissue**

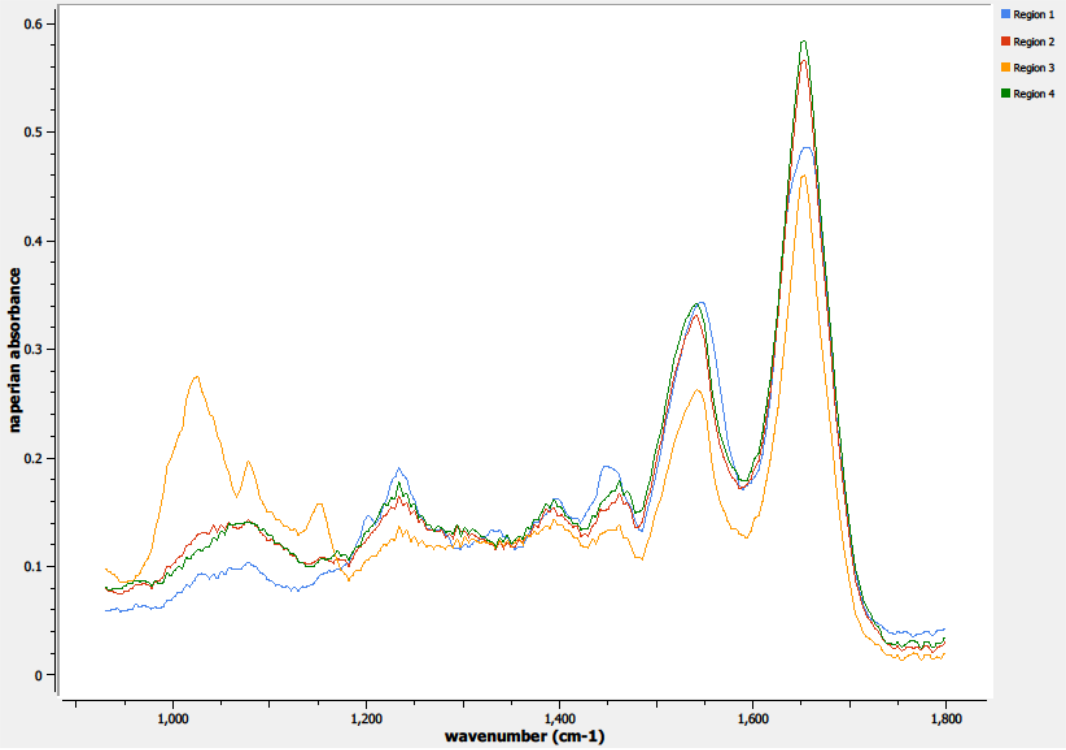



**Figure 0.22. Point P Spectra**

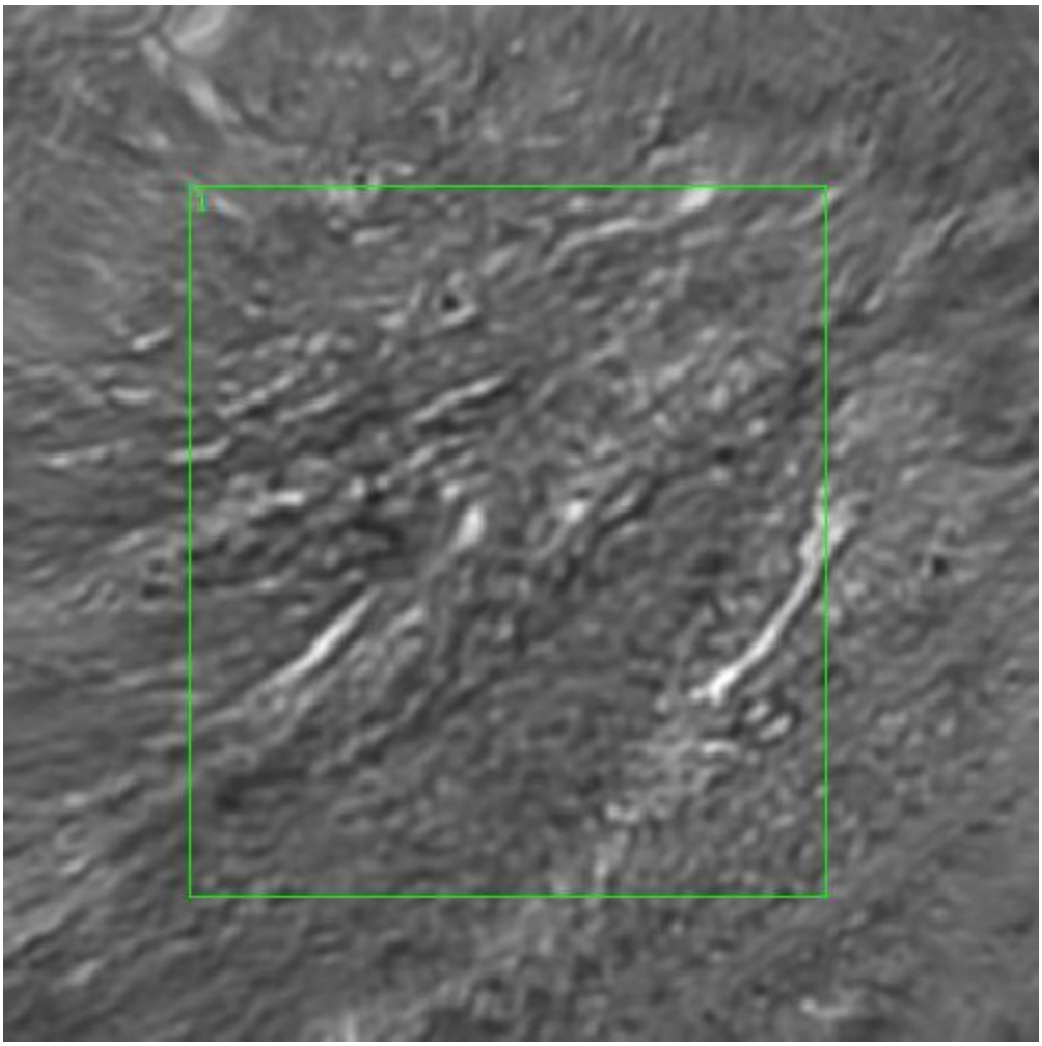


**Figure 0.23. Point Q Regions of Interest**
**Region 1: lamina propia, transition region between squamous cell carcinoma and normal tissue**

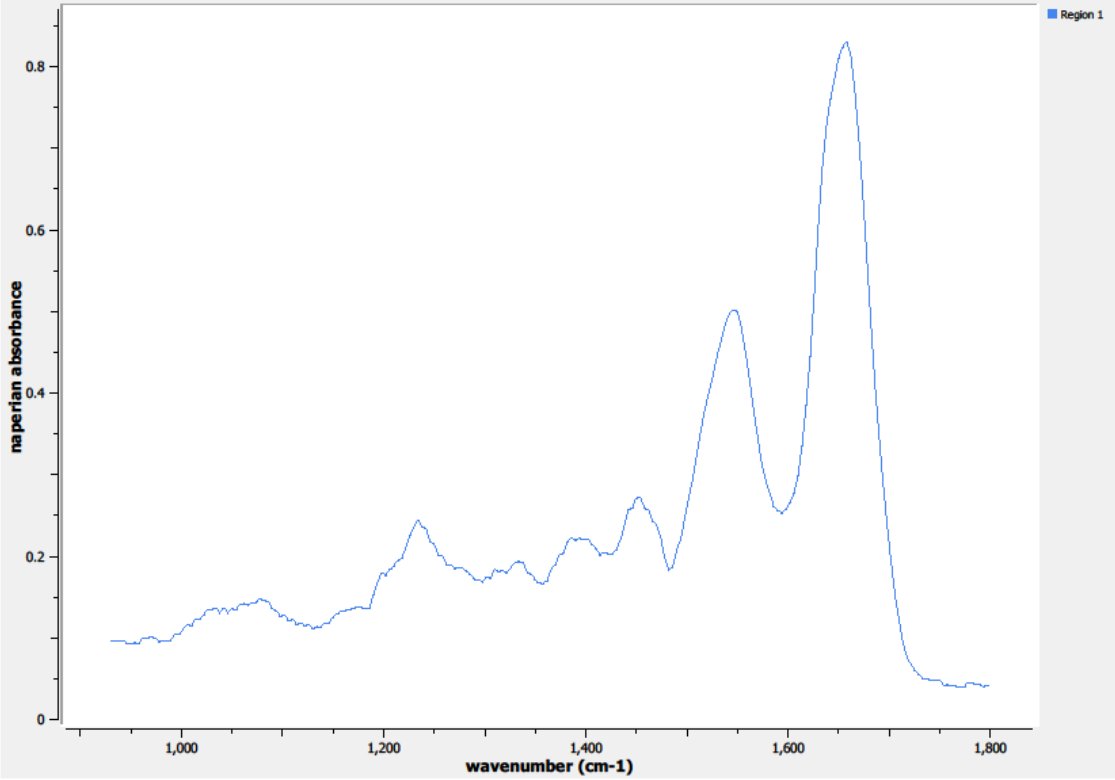



**Figure 0.24. Point Q Spectra**

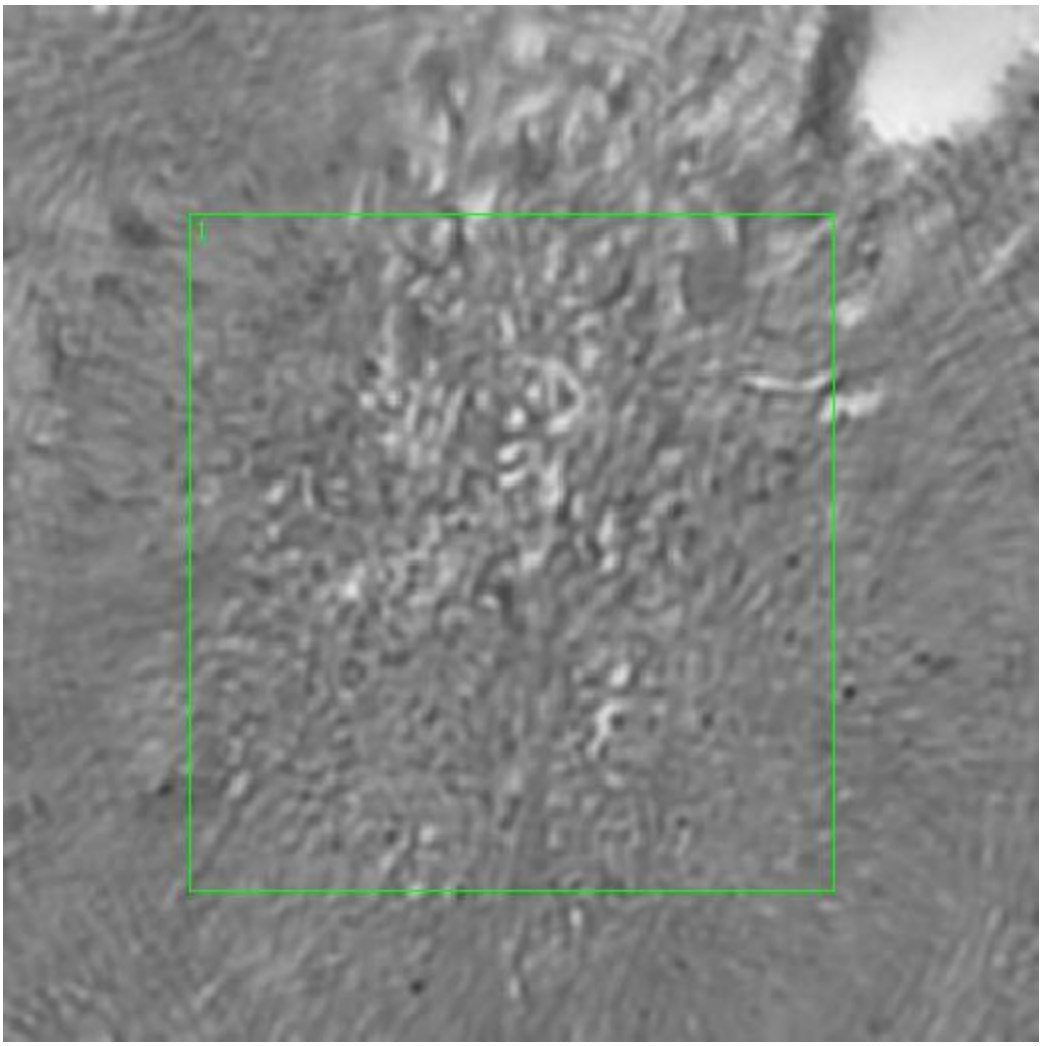


**Figure 0.25. Point R Regions of Interest**
**Region 1: lamina propia, transition region between squamous cell carcinoma and normal tissue**

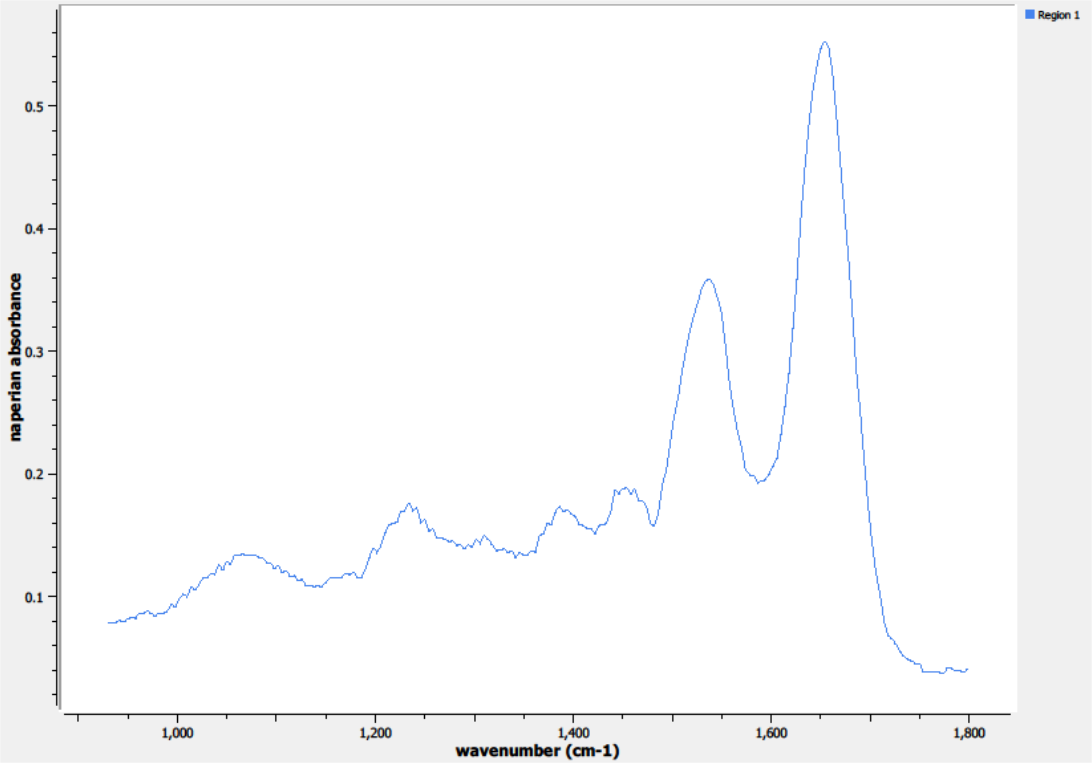



**Figure 0.26. Point R Spectra**

## A3.03-04-A370a Sample I

| | |
|---|---|
| Unique ID | 03-04-A370a Sample I |
| Age | 34 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | squamous cell carcinoma, adjacent normal ectocervix, transition region |

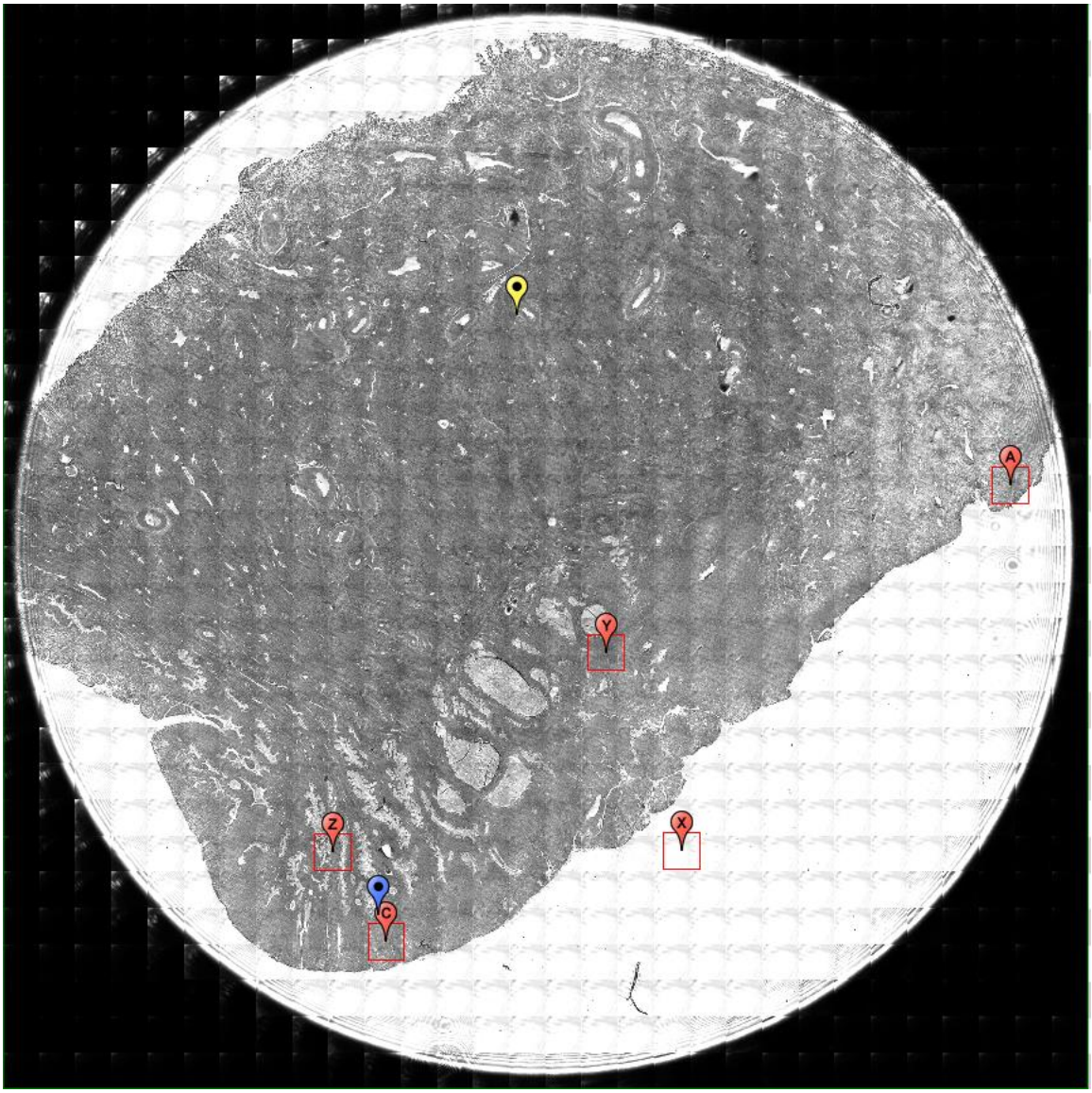



**Figure 0.27. Specimen 03-04-A370a Sample I, cervical cancer tissue**

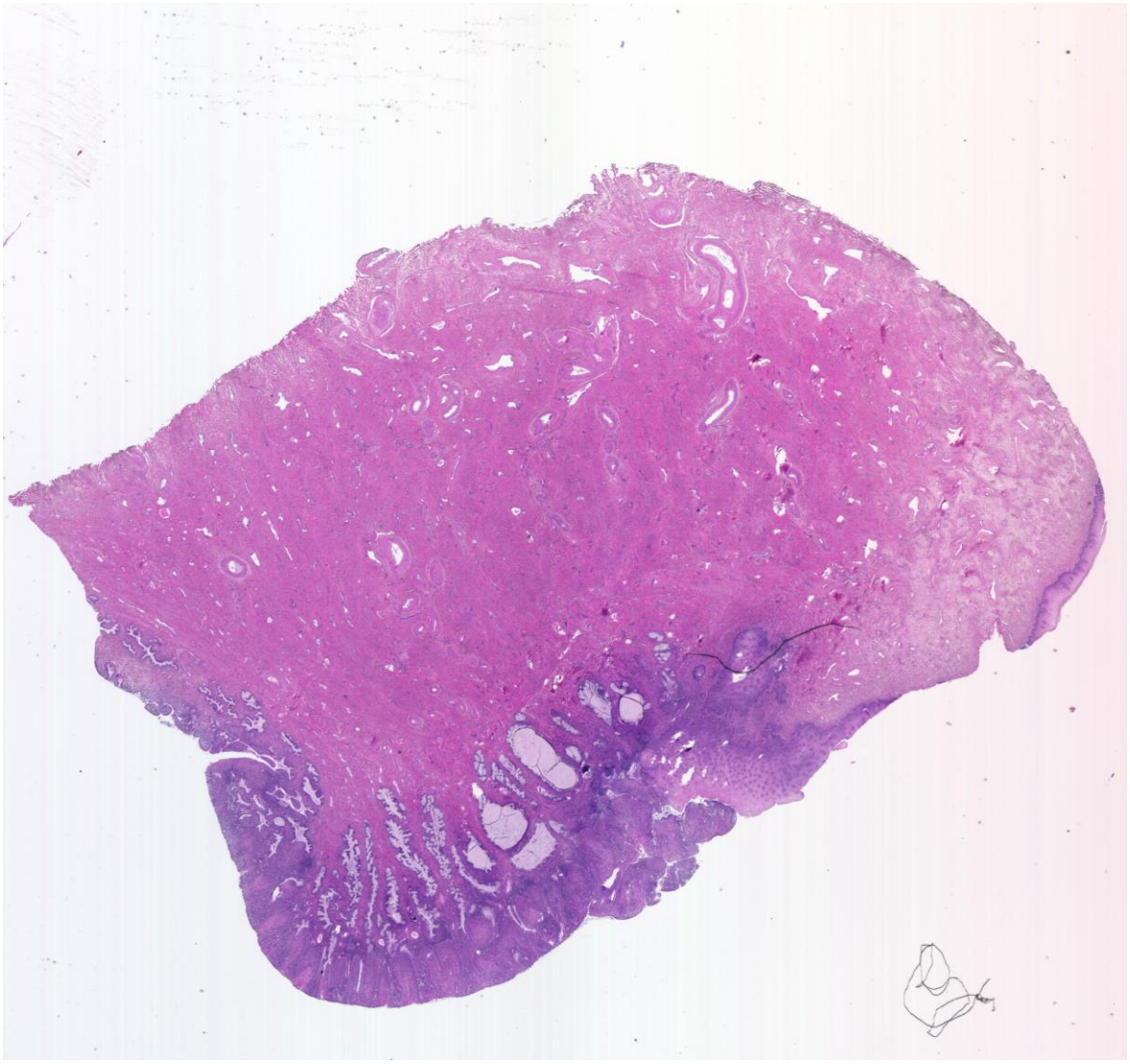


**Figure 0.28. Specimen 03-04-A370a Sample I, H&E stain, squamous cell carcinoma of the cervix**

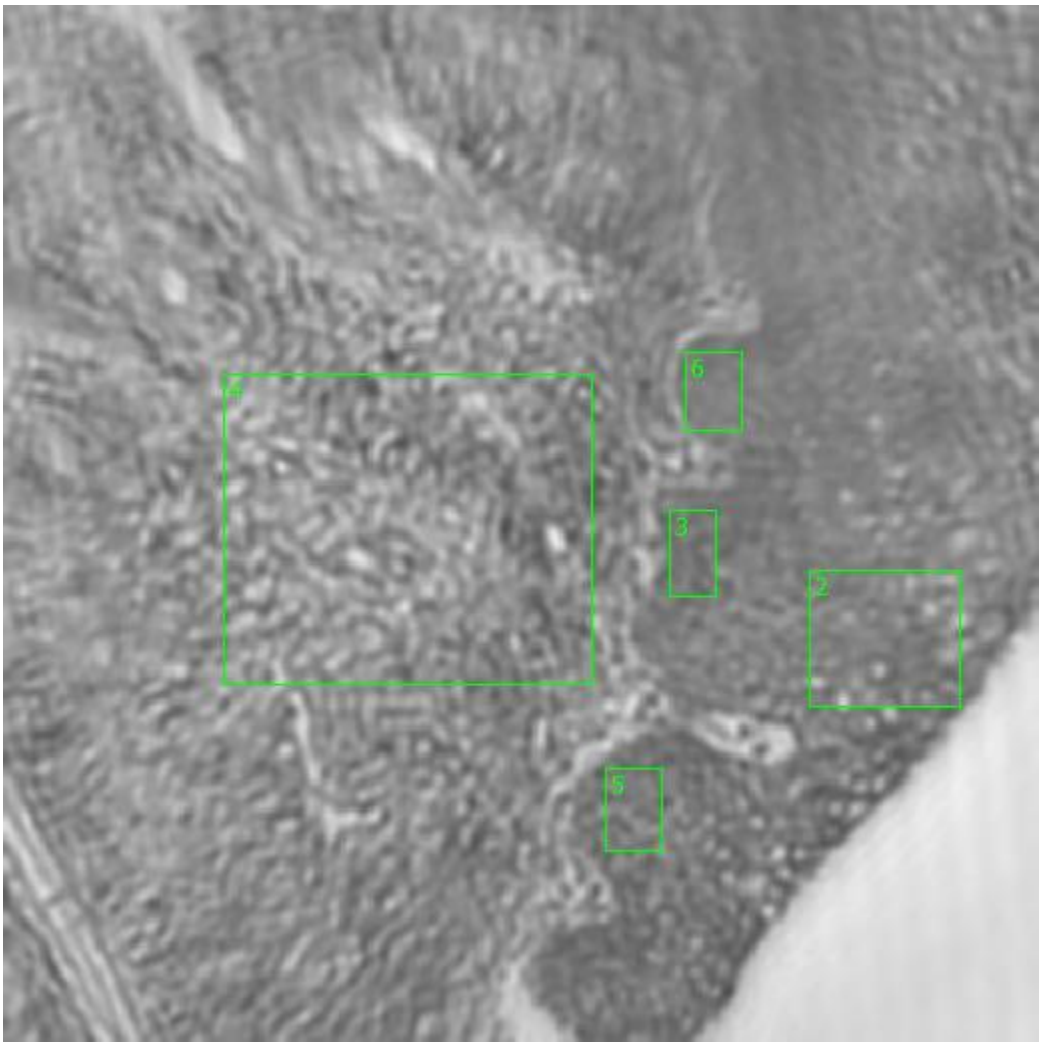



**Figure 0.29. Point A Regions of Interest**
**Region 2: columnar epithelium; transition region between squamous cell carcinoma and normal tissue**
**Regions 3, 5, 6: basal layer, parabasal layer; transition region between squamous cell carcinoma and normal tissue**
**Region 4: basement membrane, lamina propia; transition region between squamous cell carcinoma and normal tissue**

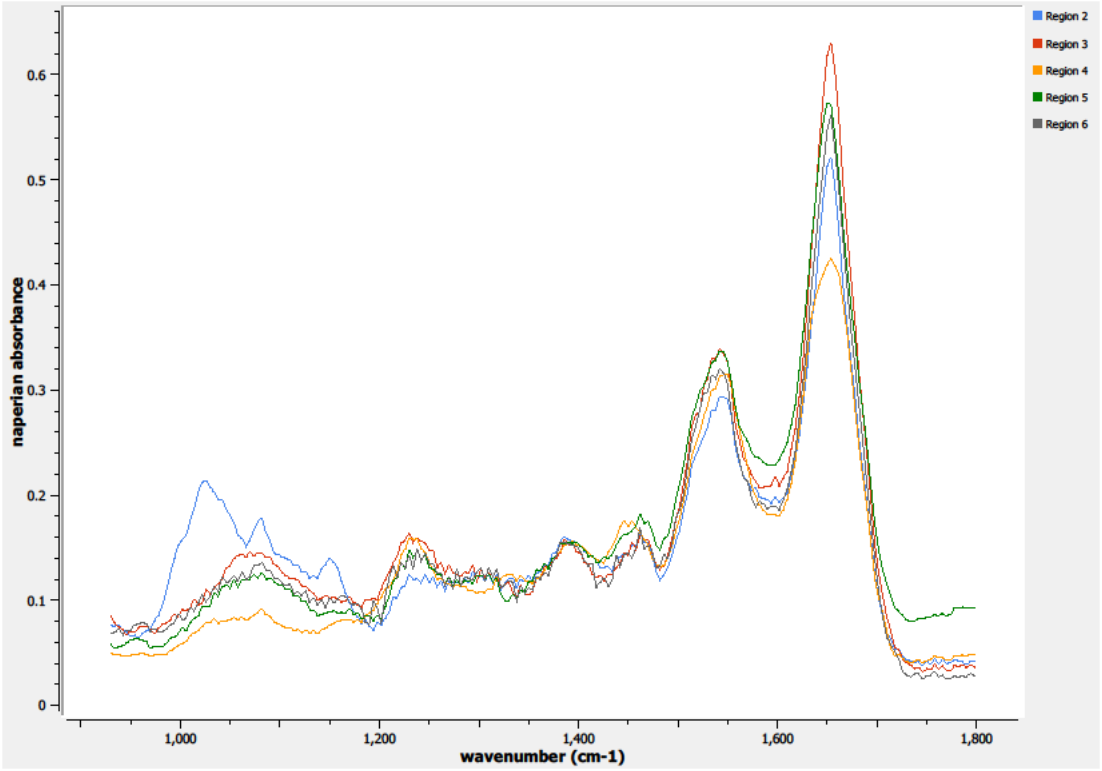



**Figure 0.30. Point A Spectra**

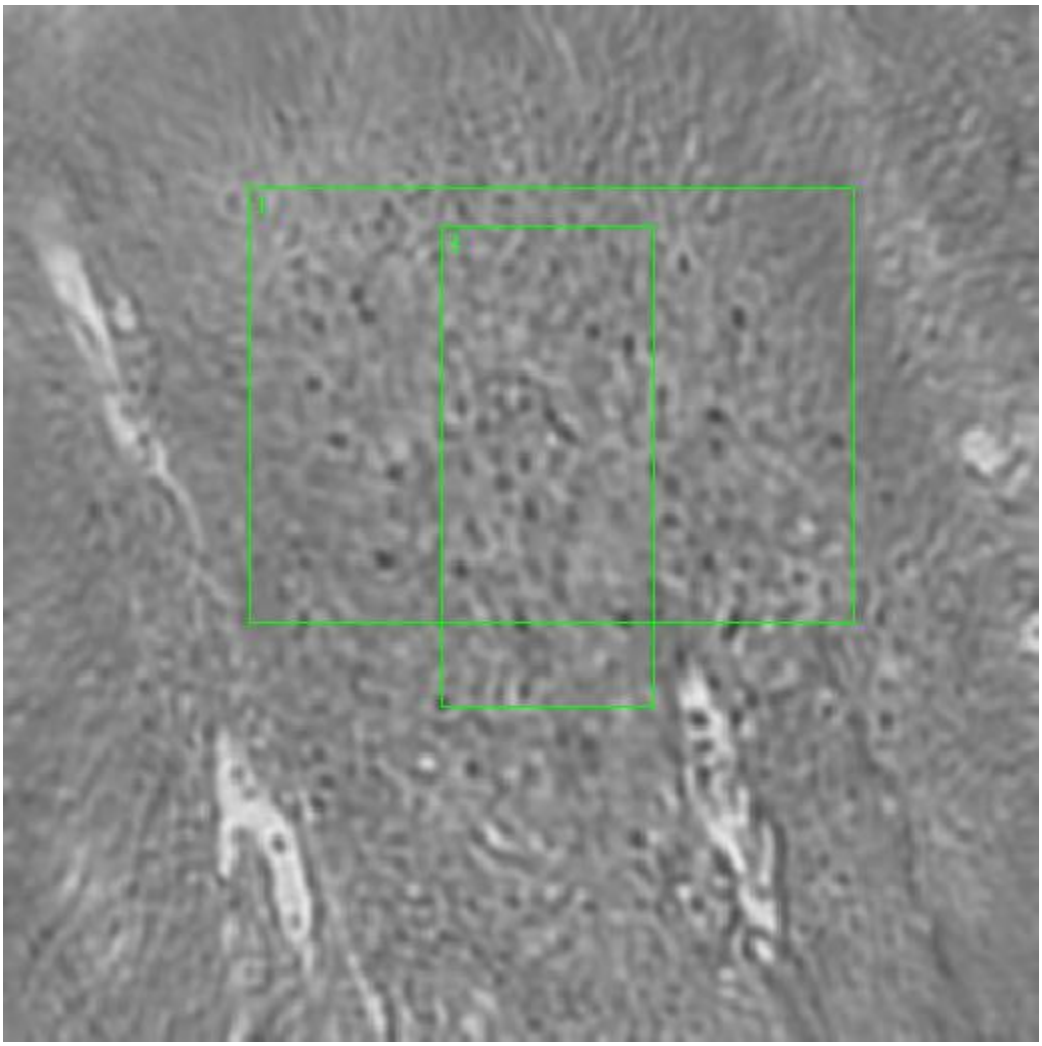


**Figure 0.31. Point C Regions of Interest**
**Regions 1, 3: columnar epithelium, squamous cell carcinoma**

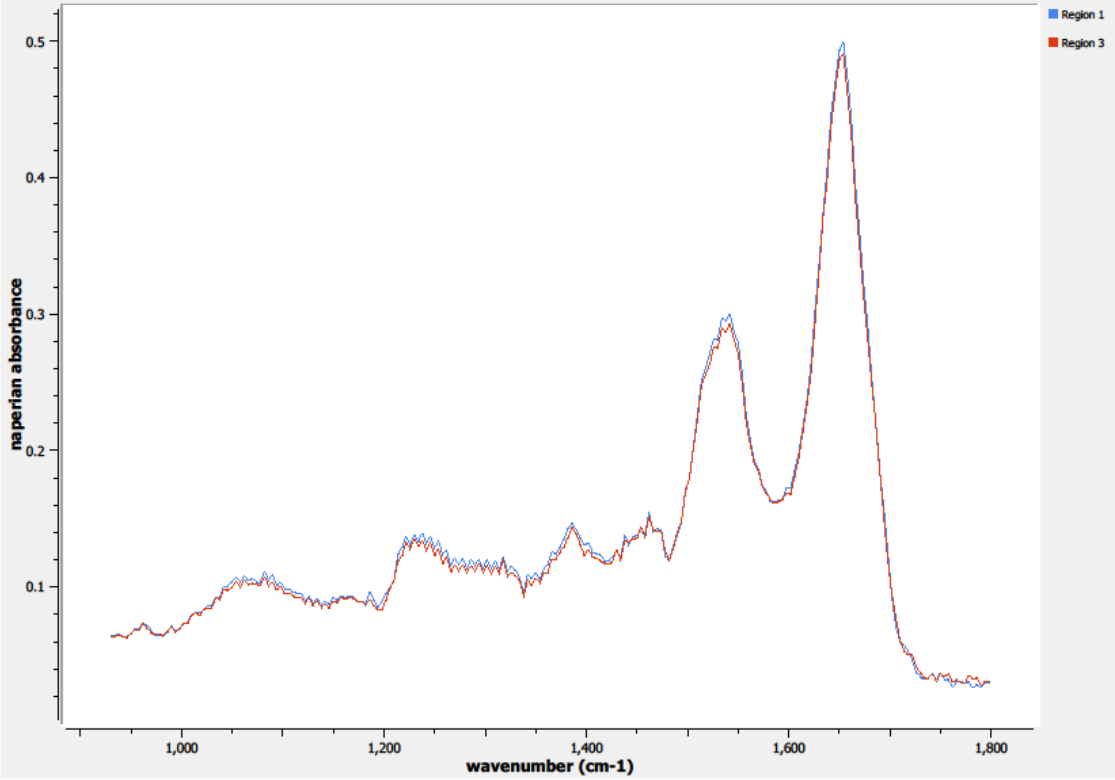



**Figure 0.32. Point C Spectra**

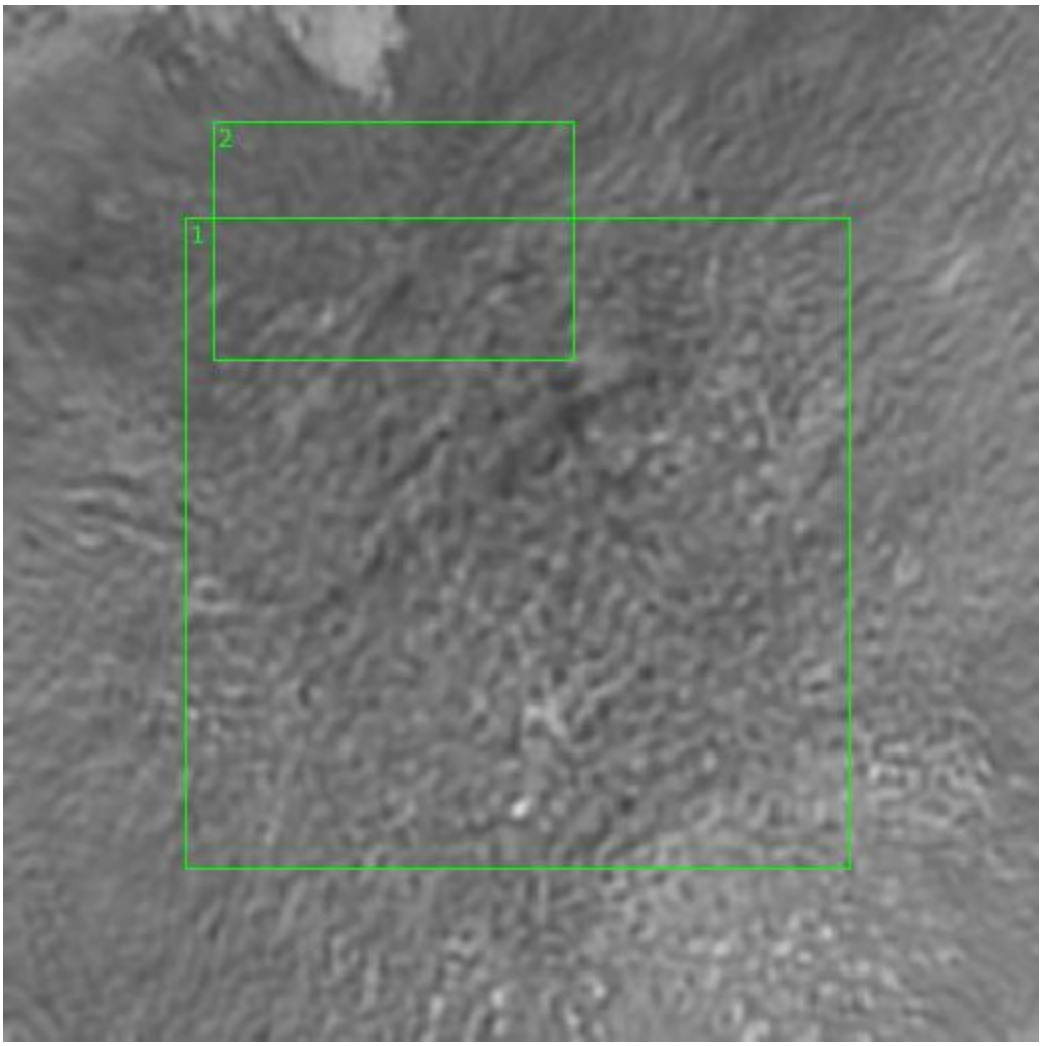



**Figure 0.33. Point Y Regions of Interest**
**Regions 1, 2: lamina propia near glandular structure, squamous cell carcinoma**

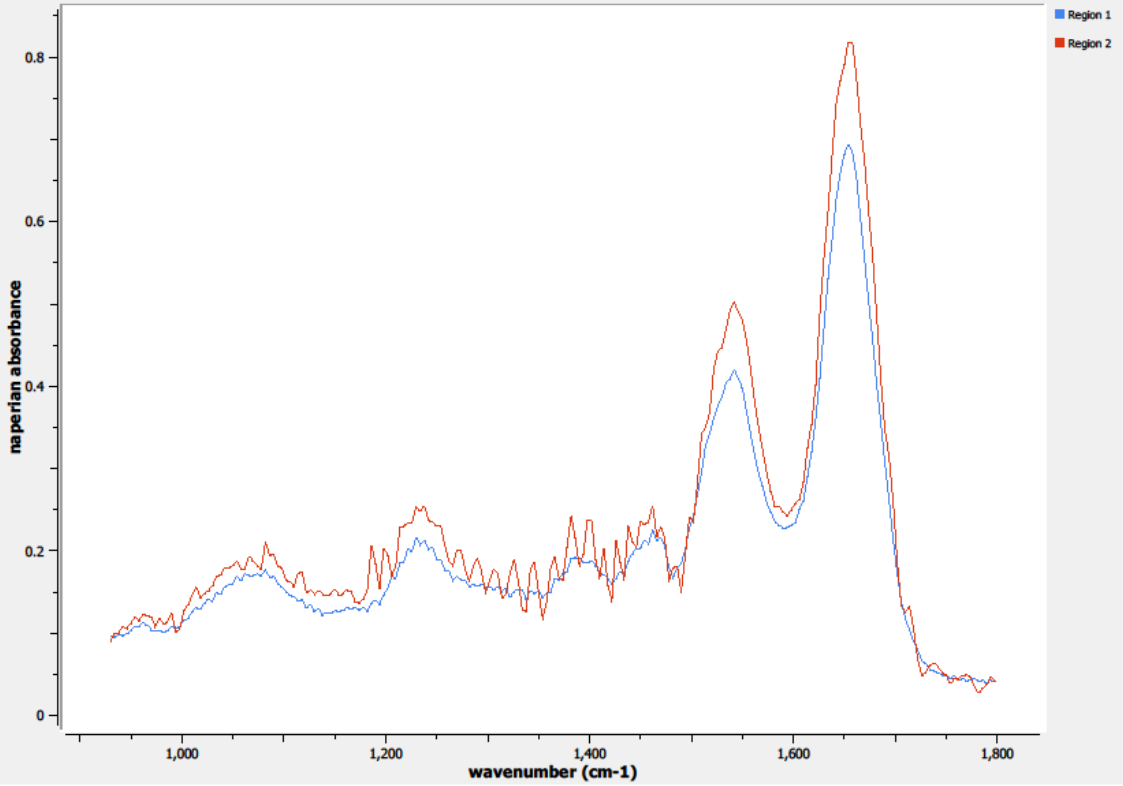



**Figure 0.34. Point Y Spectra**

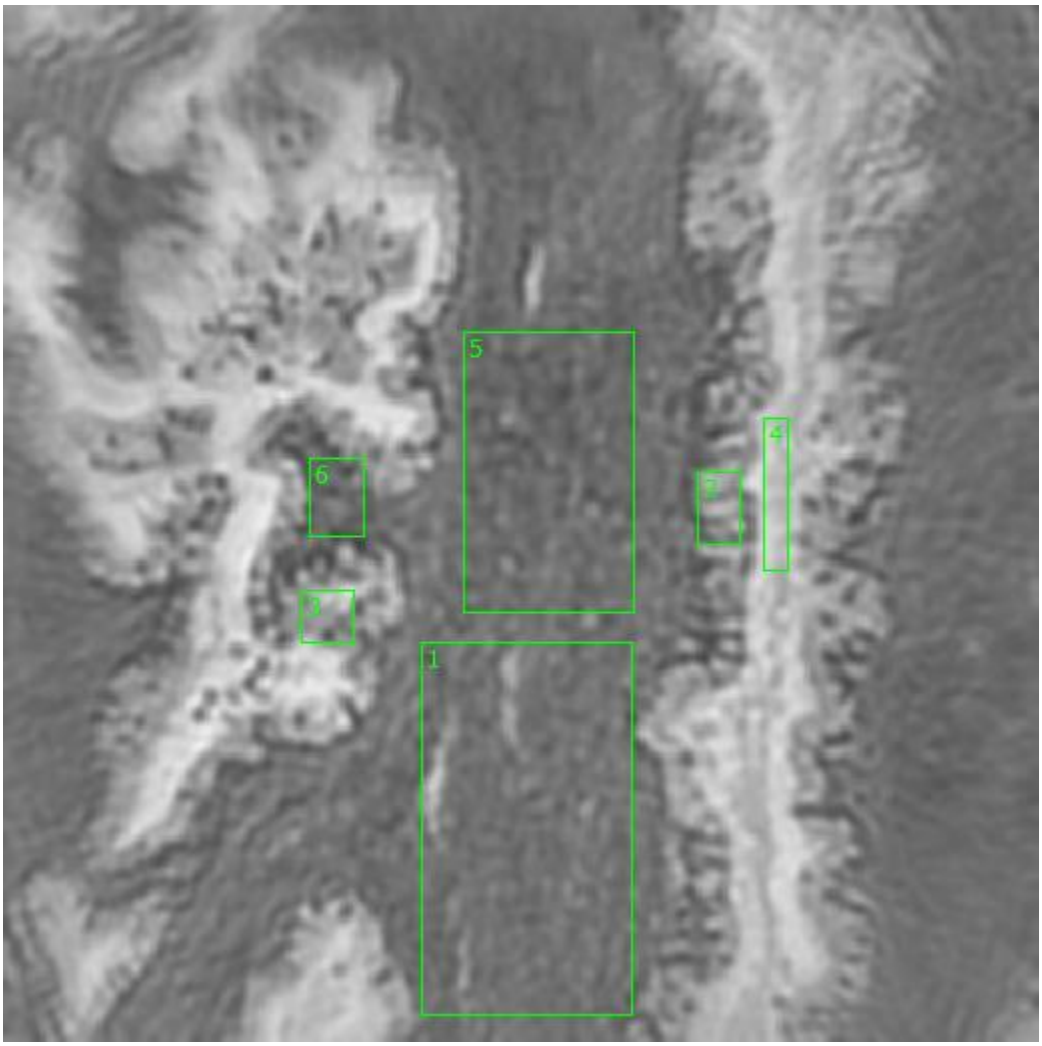



**Figure 0.35. Point Z Regions of Interest**
**Regions 1, 5, 6: lamina propia near glandular structure, squamous cell carcinoma**
**Regions 2, 3: edge of glandular structure, squamous cell carcinoma**
**Region 4: glandular structure, squamous cell carcinoma**

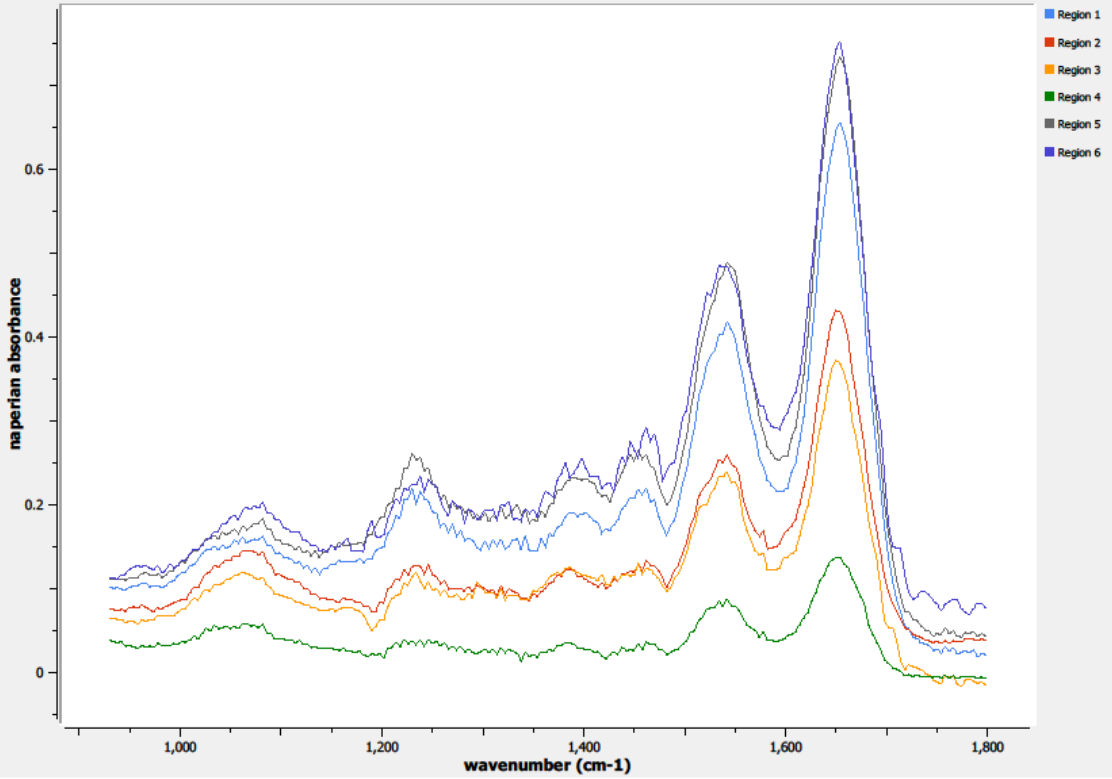



**Figure 0.36. Point Z Spectra**

## A4.03-04-A370a Sample II

| | |
|---|---|
| Unique ID | 03-04-A370a Sample II |
| Age | 34 |
| Race | white |
| Location of Sample Collection | ectocervix |
| Notes of Interest | squamous cell carcinoma, adjacent normal ectocervix, transition region |

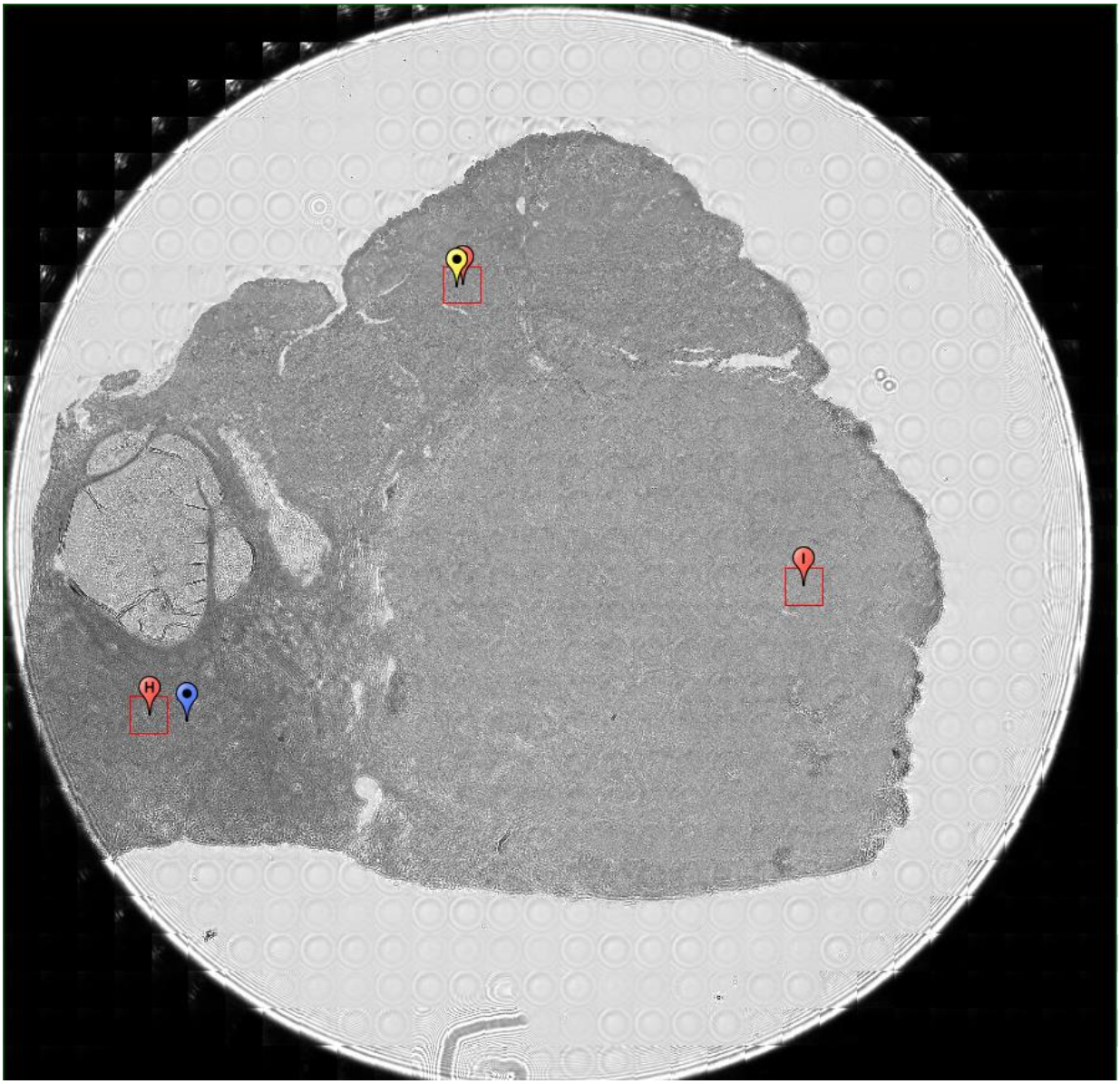


**Figure 0.37. Specimen 03-04-A370a Sample II, cervical cancer tissue**

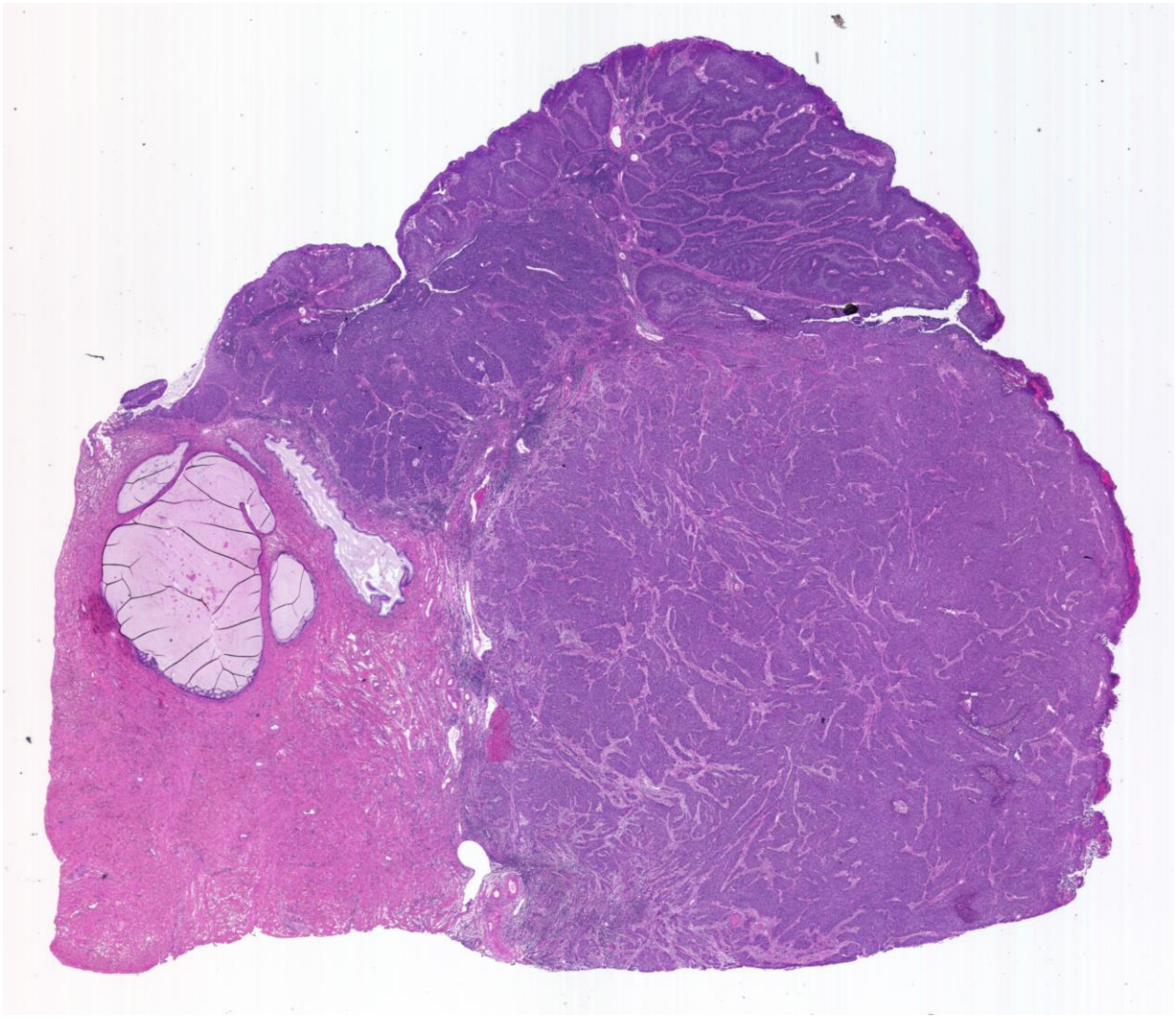


**Figure 0.38. Specimen 03-04-A370a Sample II, H&E stain, squamous cell carcinoma of the cervix**

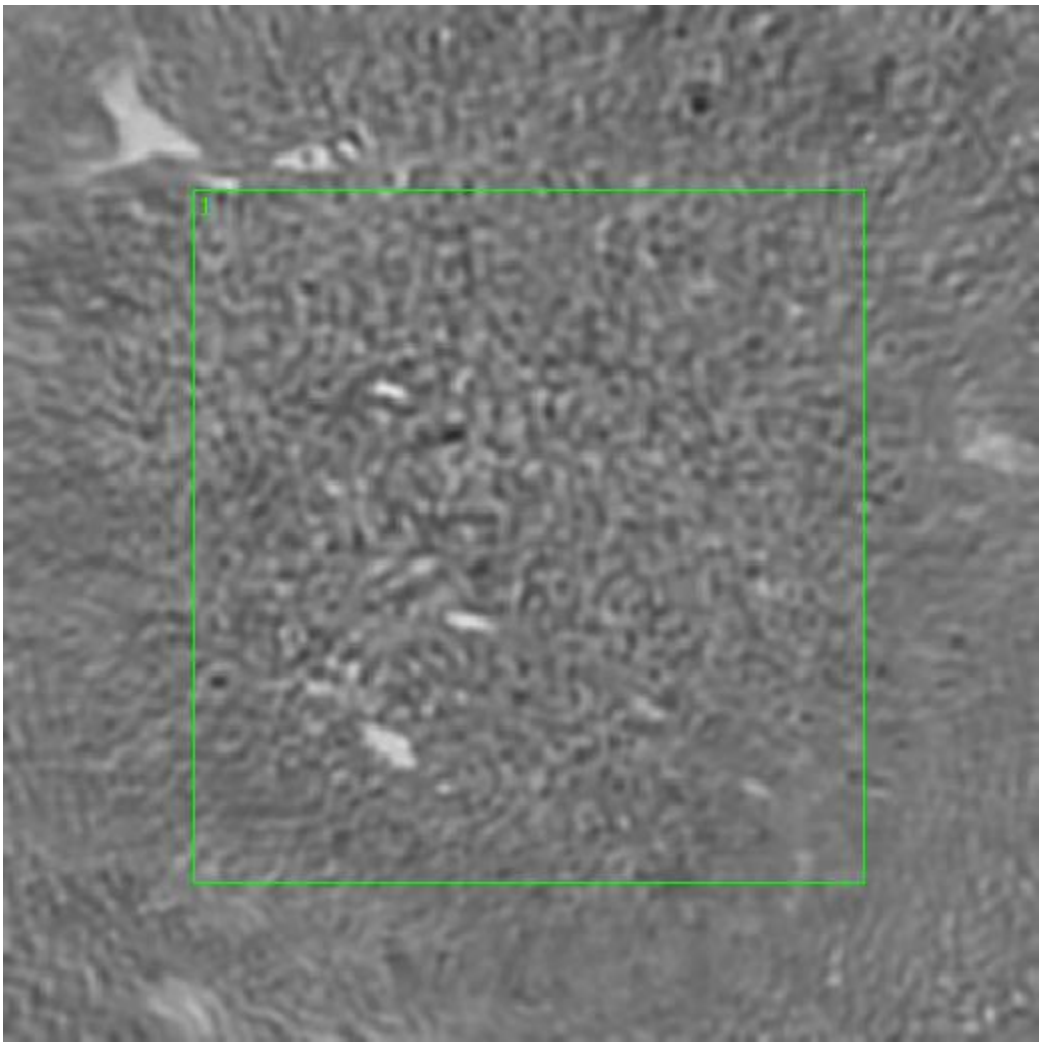


**Figure 0.39. Point B Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

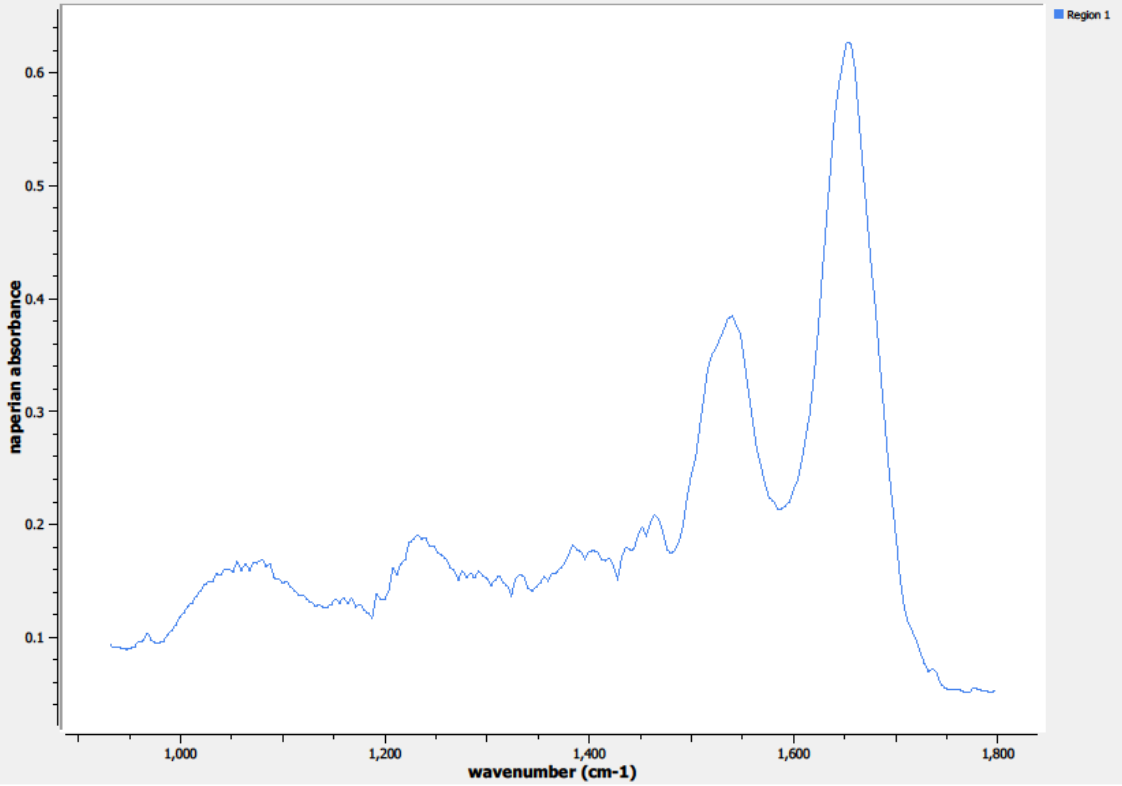



**Figure 0.40. Point B Spectra**

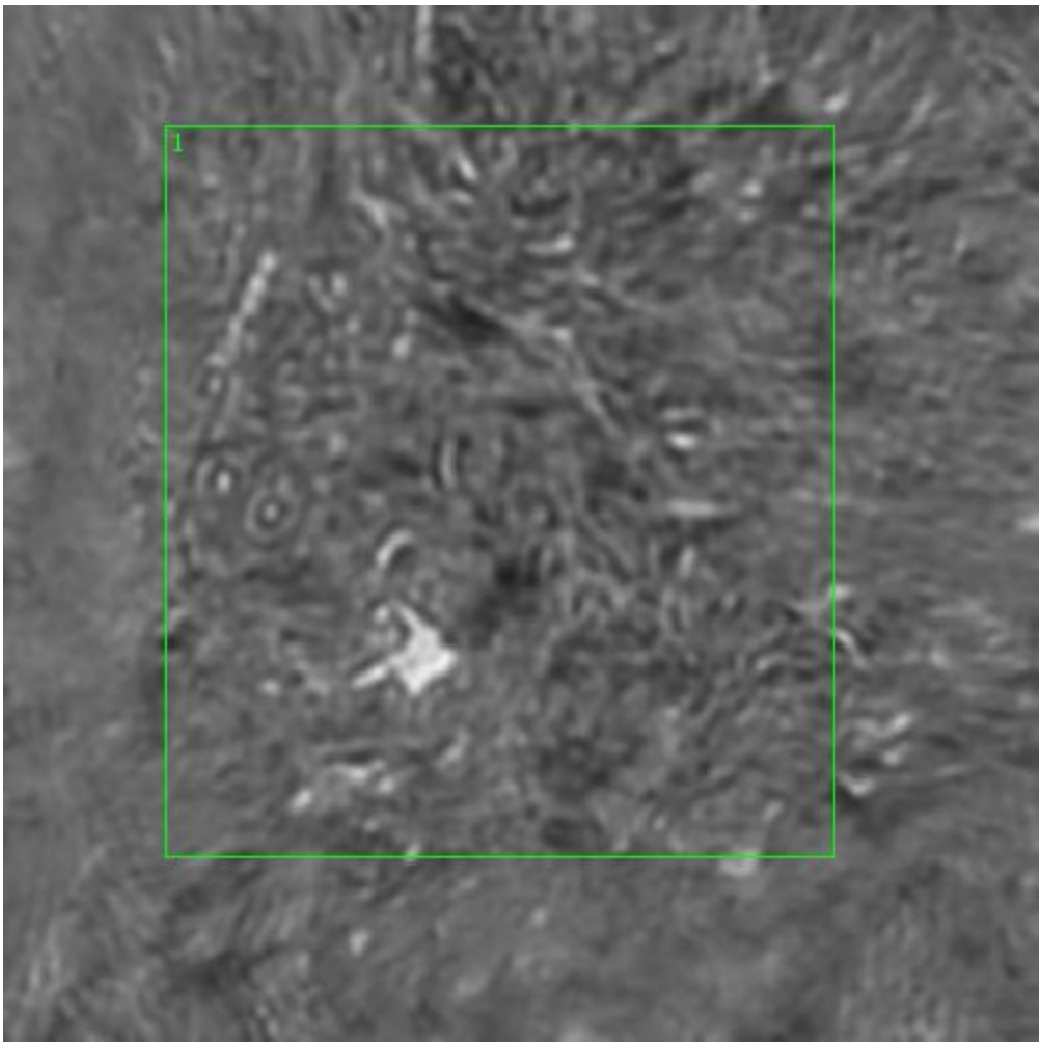



**Figure 0.41. Point H Regions of Interest**
**Region 1: lamina propia, transition region between squamous cell carcinoma and normal tissue**

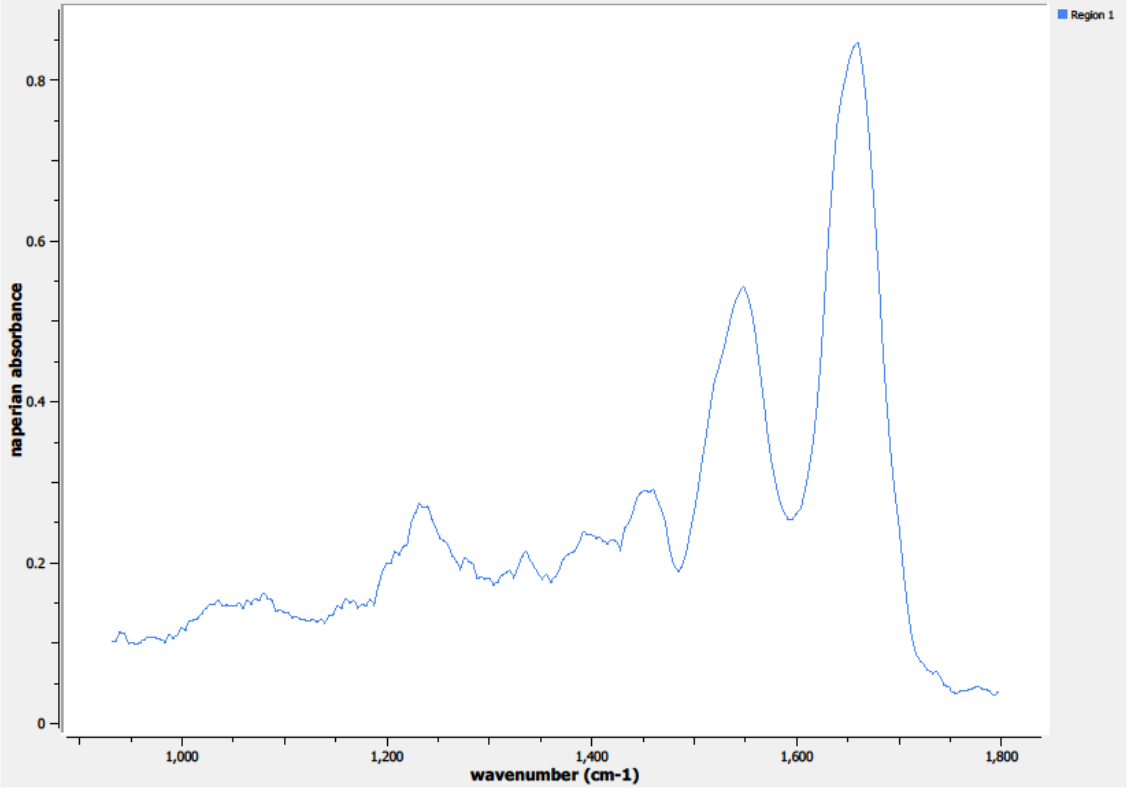



**Figure 0.42. Point H Spectra**

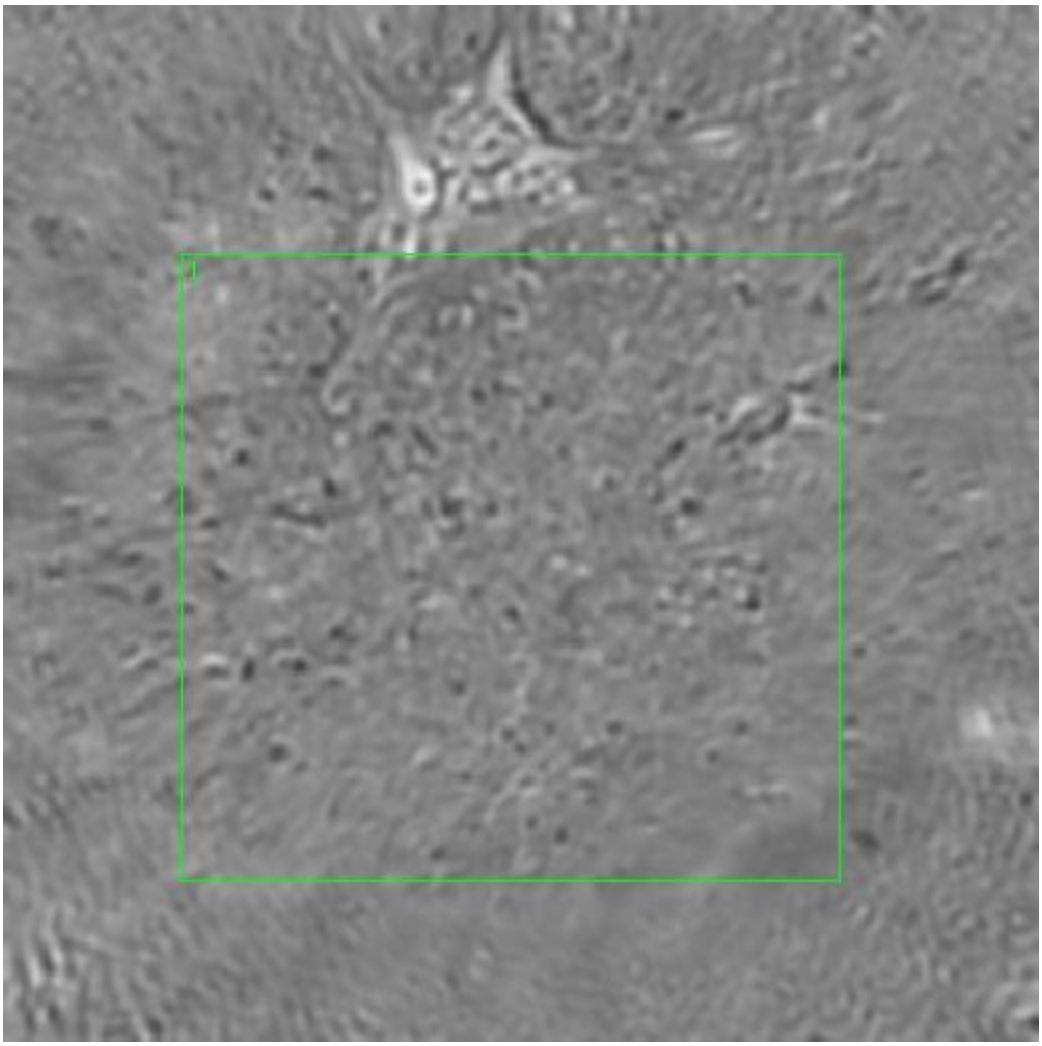


**Figure 0.43. Point I Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

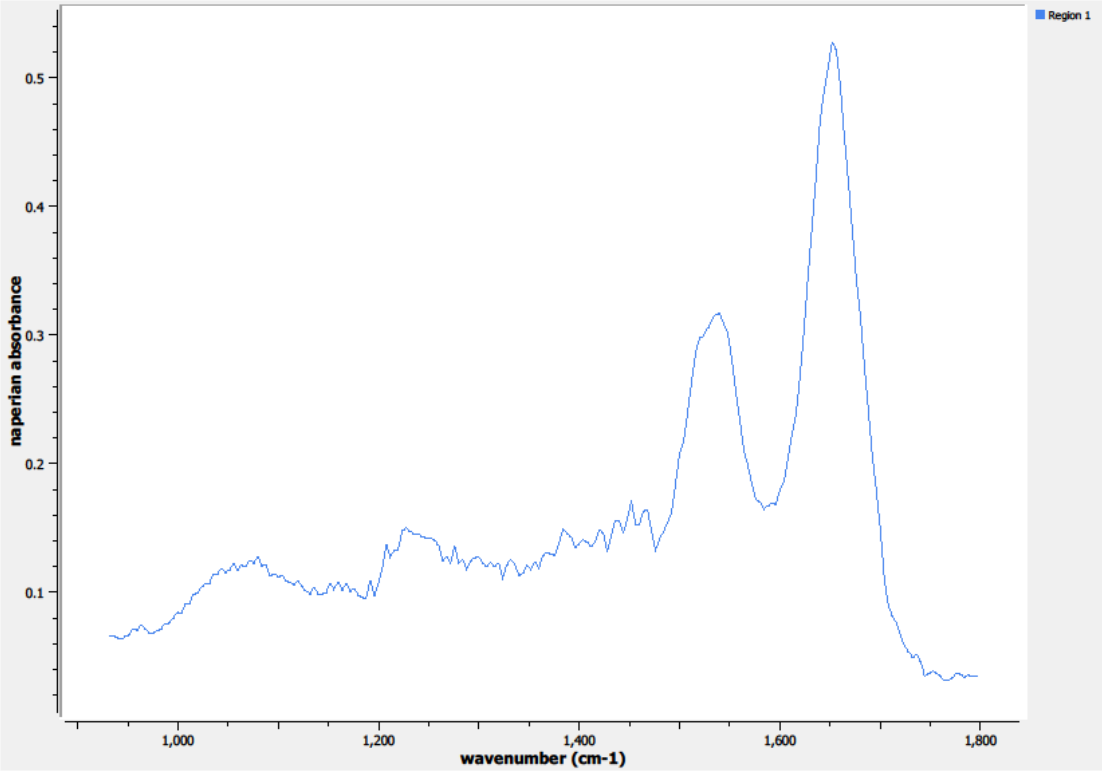



**Figure 0.44. Point I Spectra**

## A5.04-11-A257b

| | |
|---|---|
| Unique ID | 04-11-A257b |
| Age | 52 |
| Race | black |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix malignant, squamous cell carcinoma, primary, %T (80) / %N (20), data glitch at Point M, may not use |

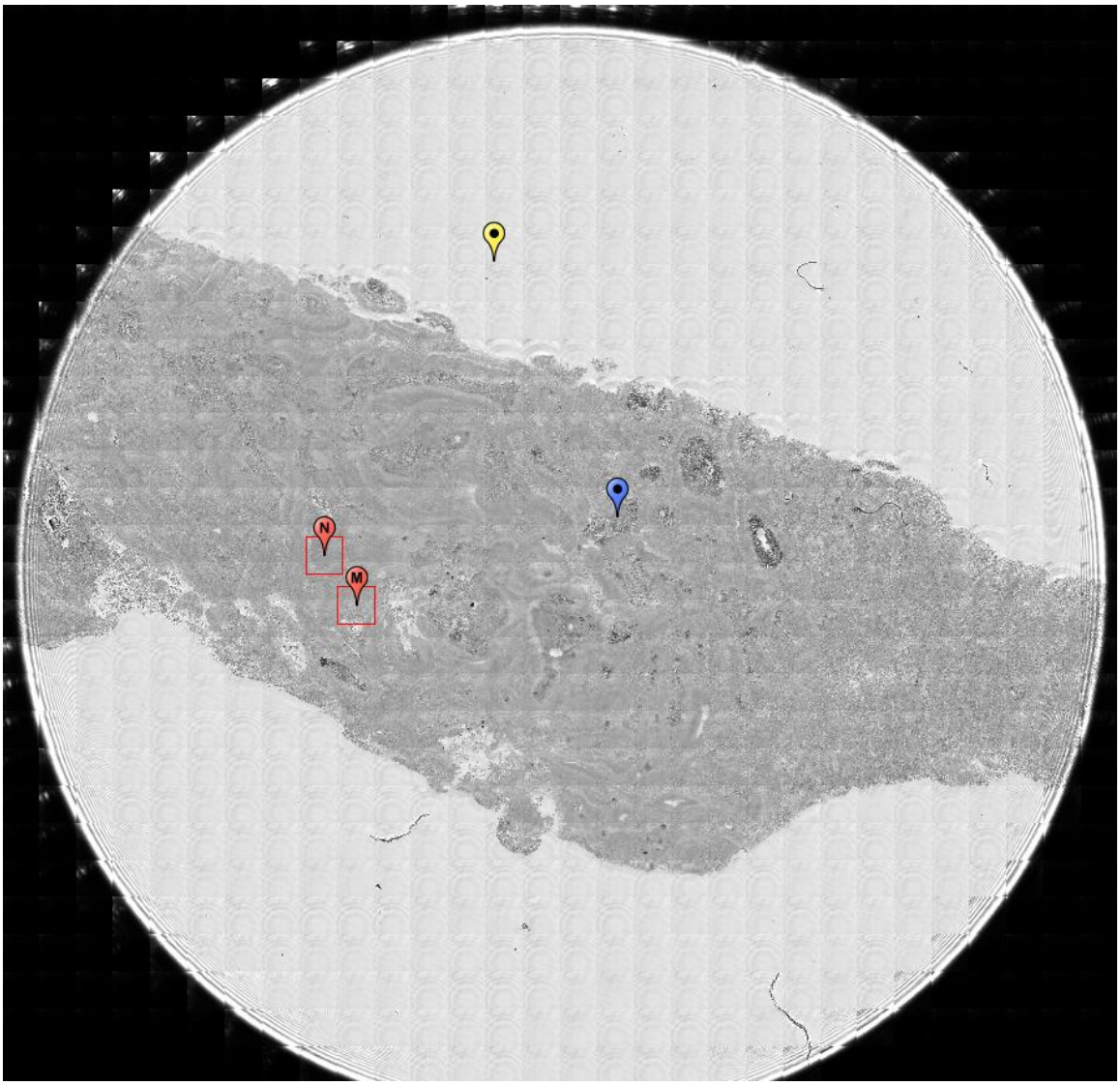



**Figure 0.45. Specimen 04-11-A257b, cervical cancer tissue**

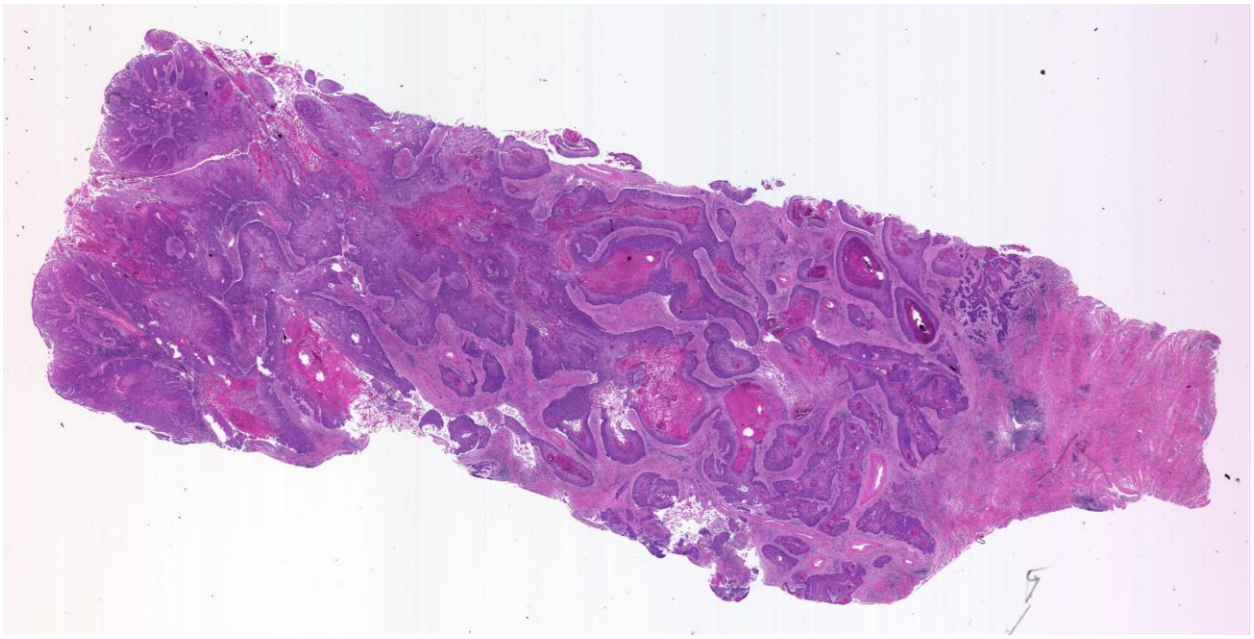


**Figure 0.46. Specimen 04-11-A257b, H&E stain, squamous cell carcinoma of the cervix**

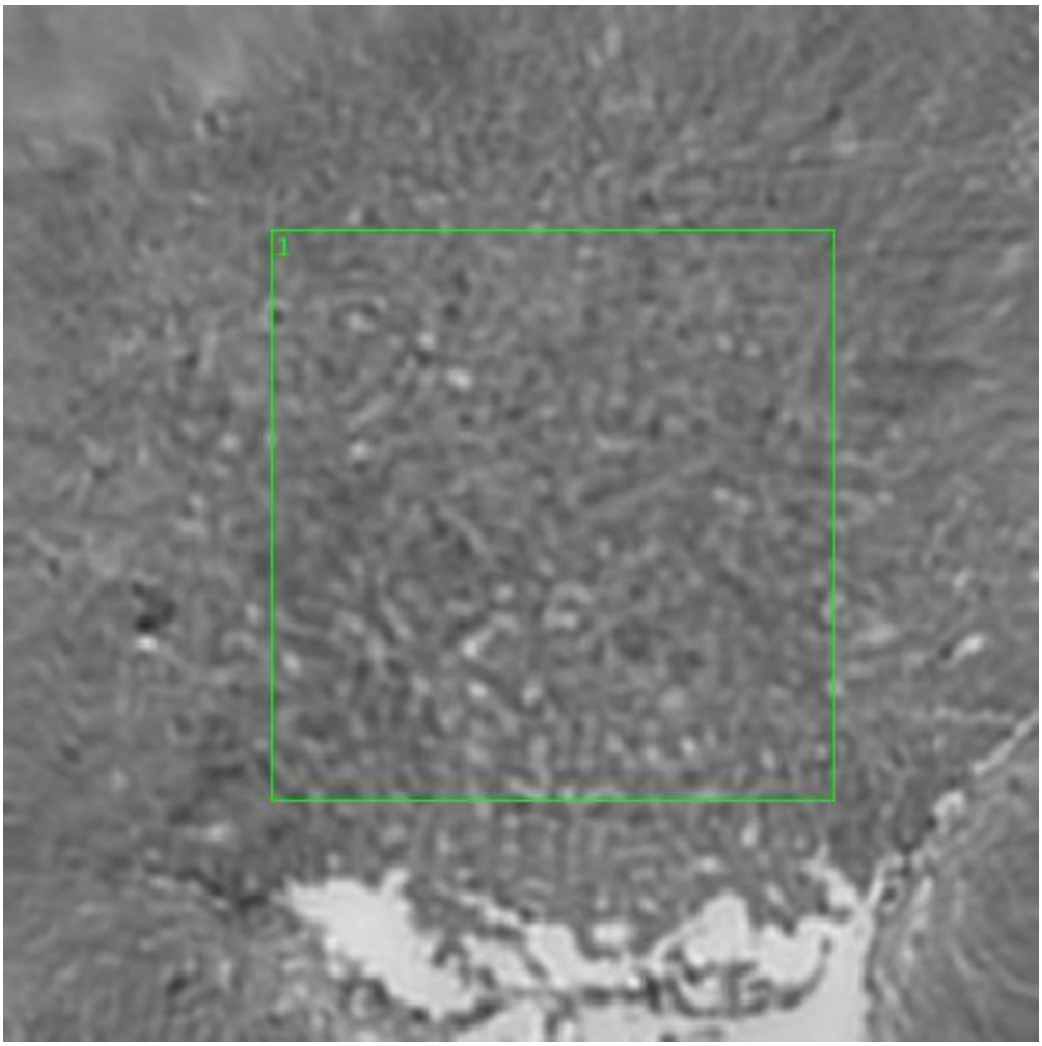



**Figure 0.47. Point M Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

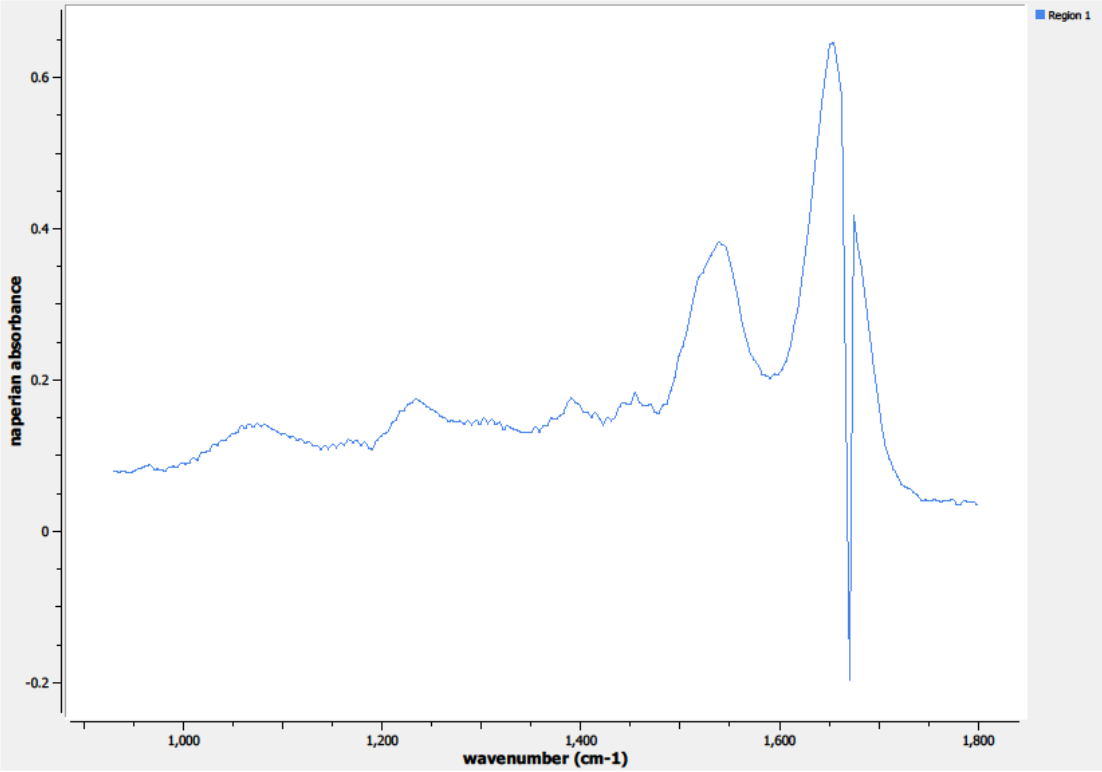



**Figure 0.48. Point M Spectra (data glitch, may not use)**

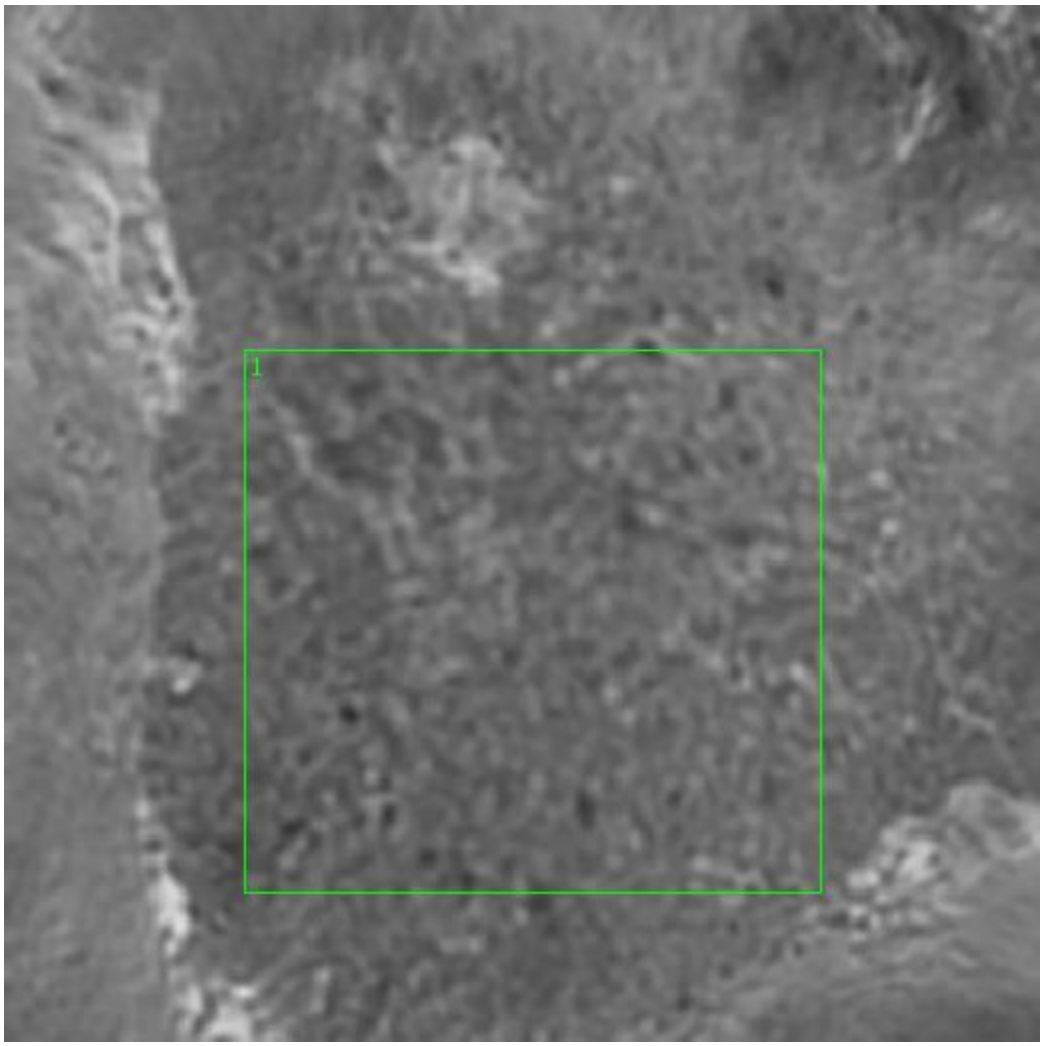



**Figure 0.49. Point N Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

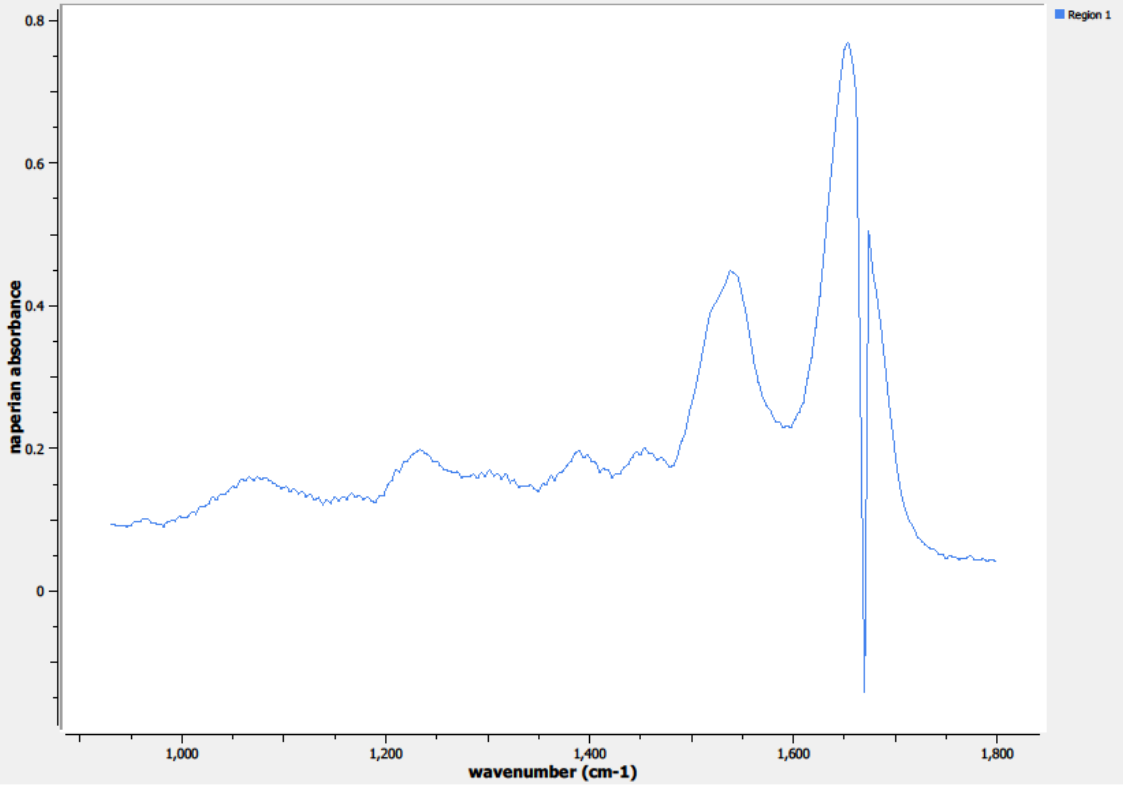



**Figure 0.50. Point N Spectra (data glitch, may not use)**

## A6.06-02-A042b

| | |
|---|---|
| Unique ID | 06-02-A042b |
| Age | 43 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix malignant, adenosquamous carcinoma, glassy cell, primary, %T (50) / %N (0) |

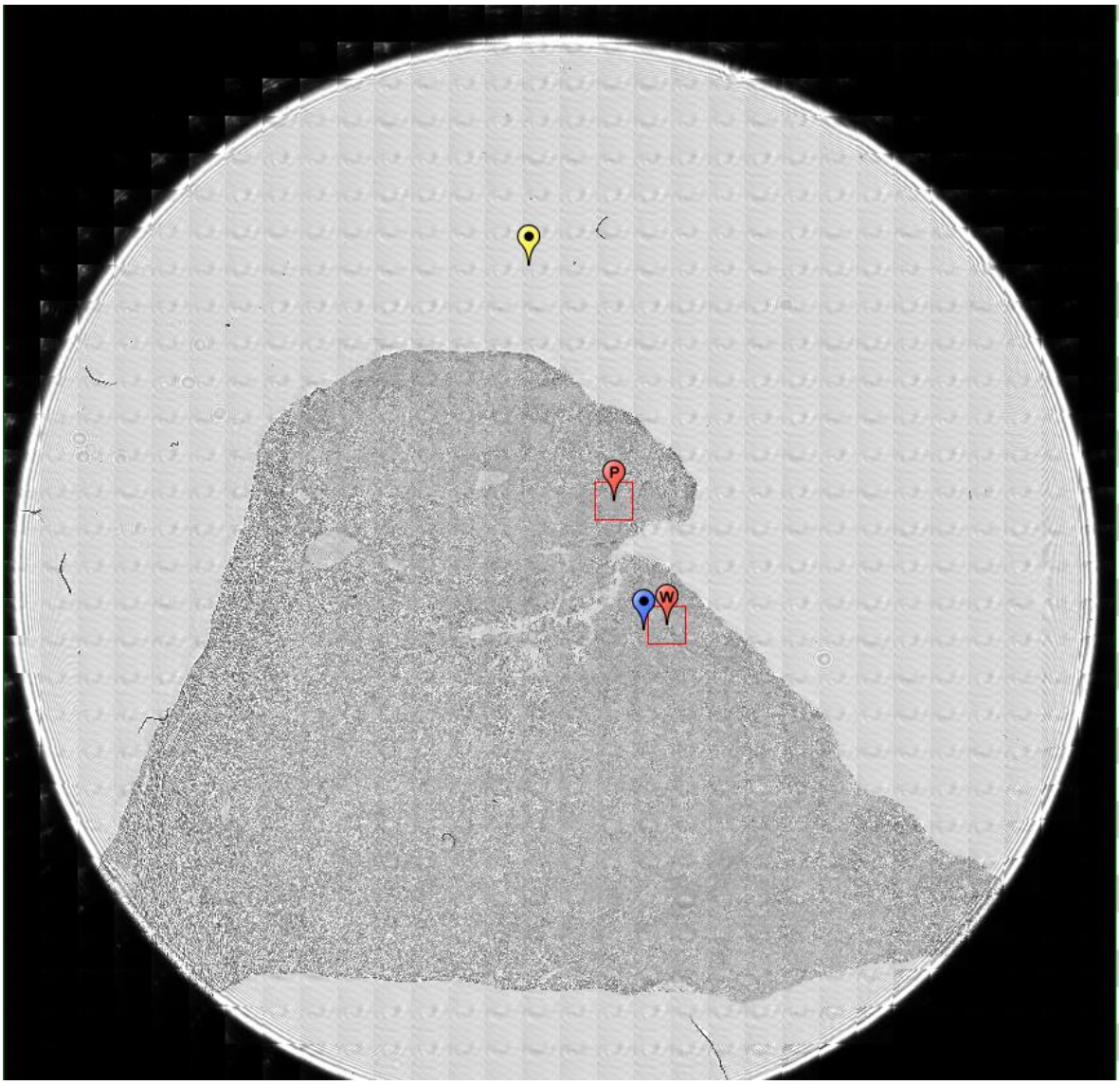



**Figure 0.51. Specimen 06-02-A042b, cervical cancer tissue**

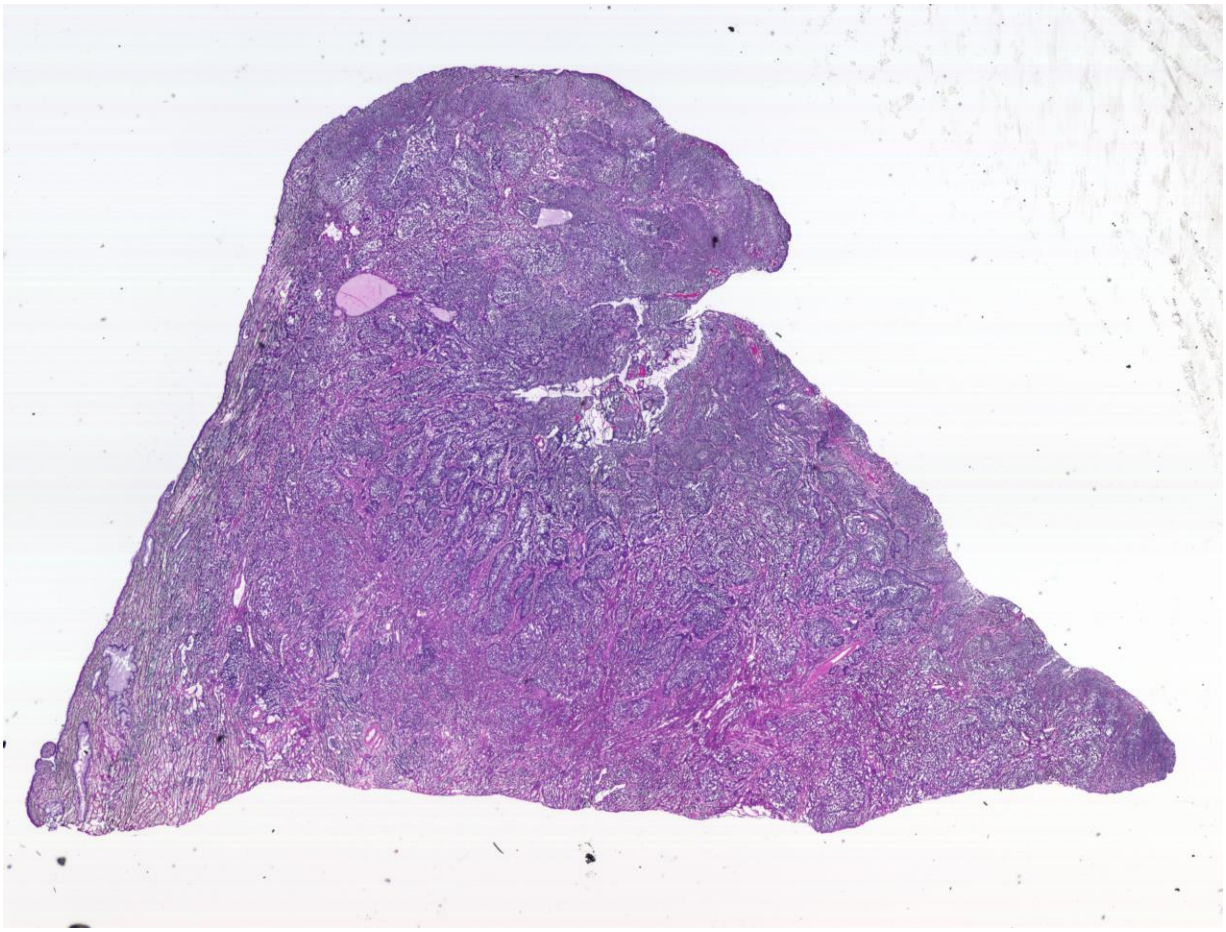


**Figure 0.52. Specimen 06-02-A042b, H&E stain, adenosquamous carcinoma of the cervix, glassy cell**

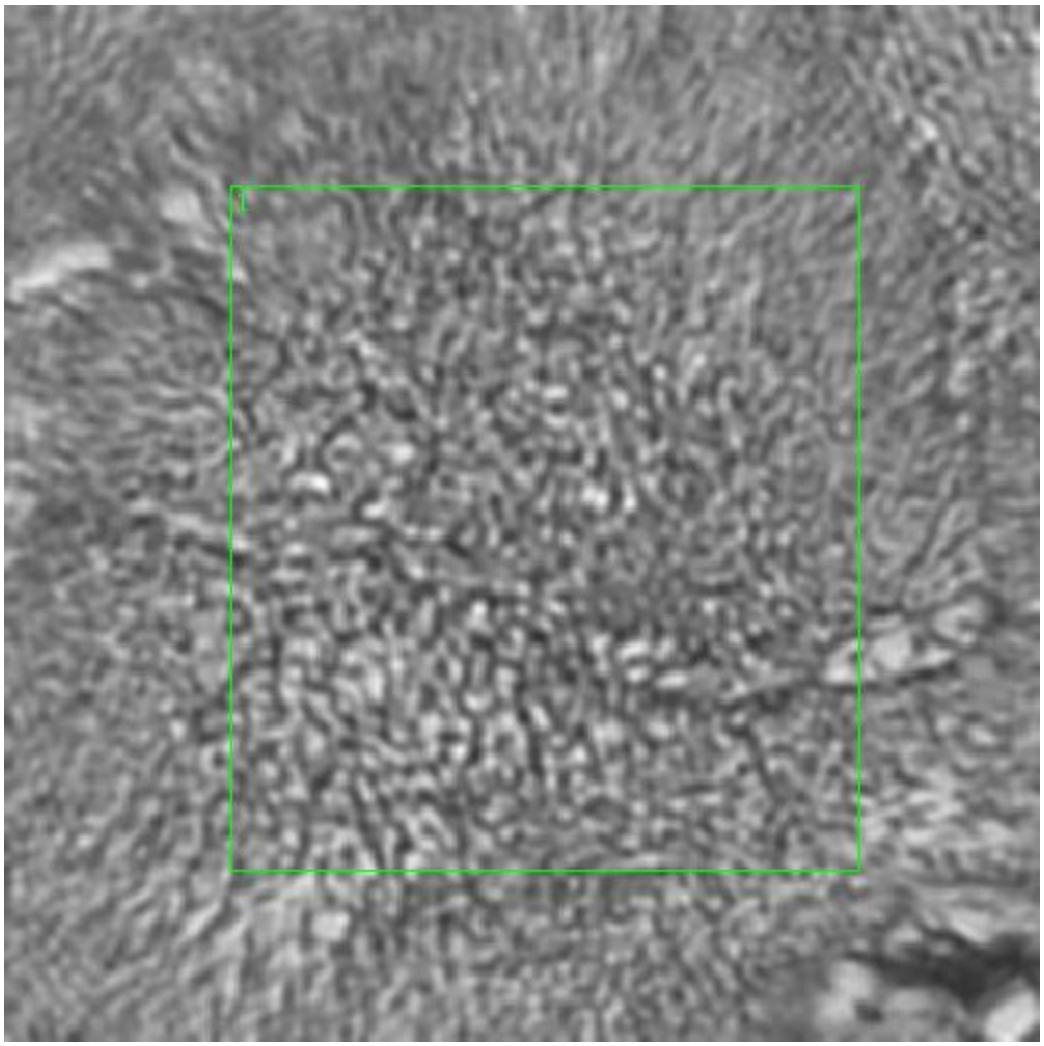


**Figure 0.53. Point P Regions of Interest**
**Region 1: lamina propia, adenosquamous glassy cell carcinoma**

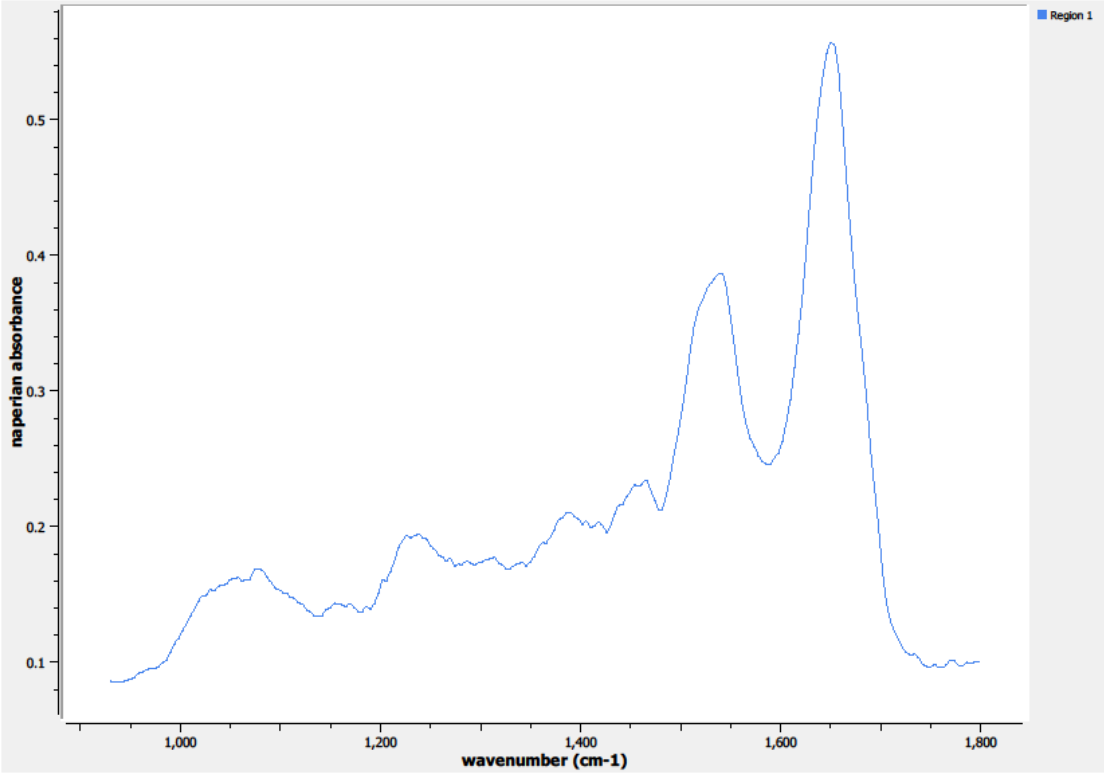



**Figure 0.54. Point P Spectra**

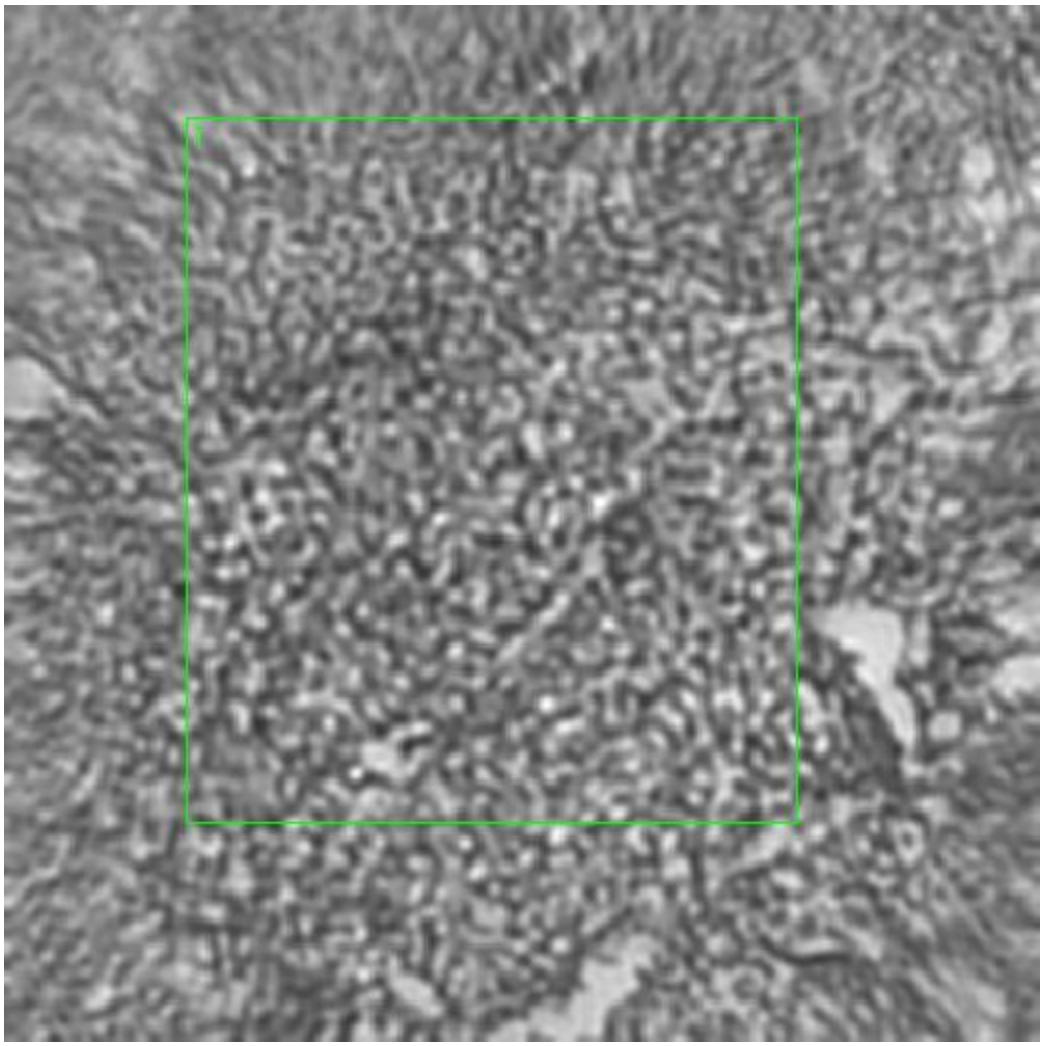


**Figure 0.55. Point W Regions of Interest**
**Region 1: lamina propia, adenosquamous glassy cell carcinoma**

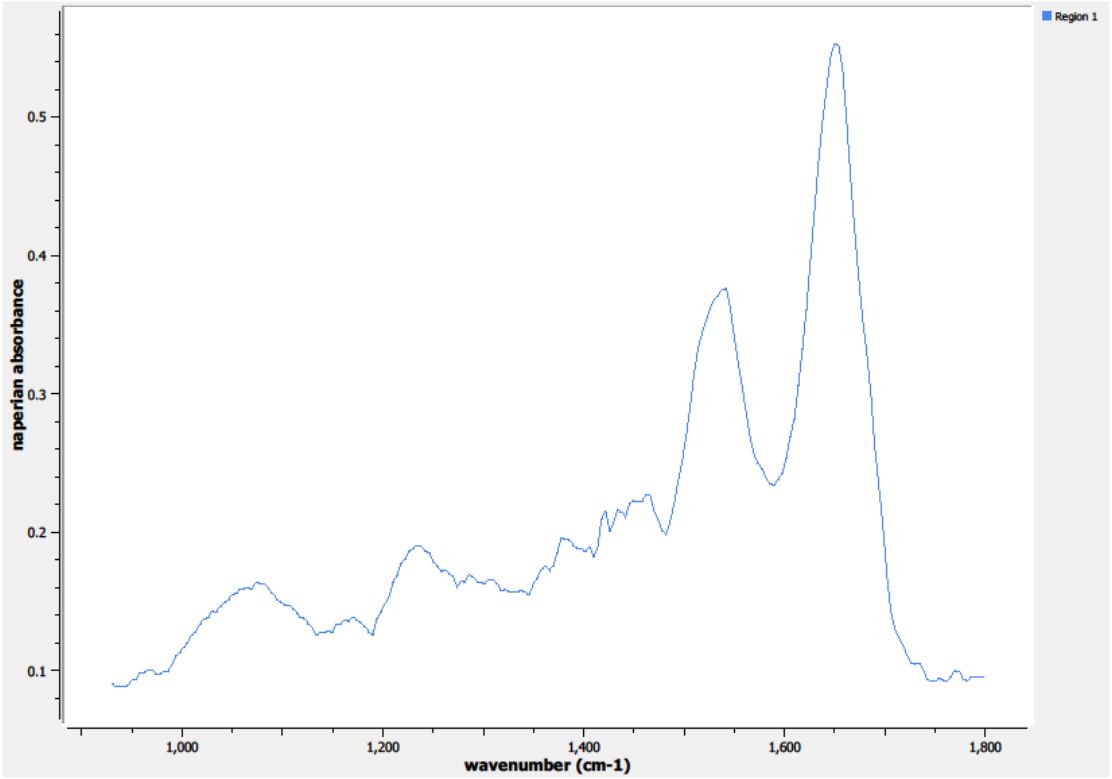



**Figure 0.56. Point W Spectra**

## A7.06-05-A276c

| | |
|---|---|
| Unique ID | 06-05-A276c |
| Age | 61 |
| Race | Asian |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix malignant, squamous cell carcinoma, primary, %T (20) / %N (10), transition region |

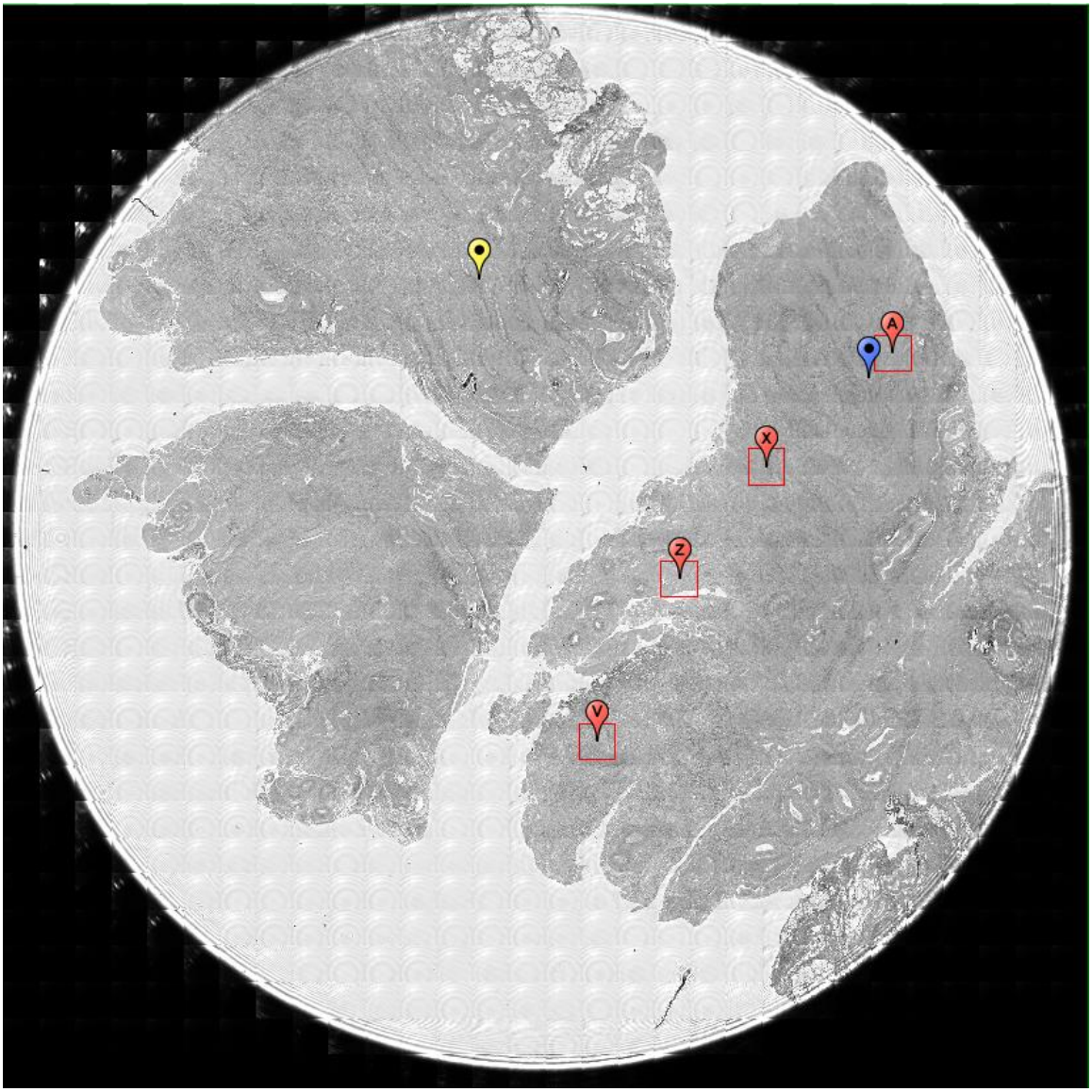



**Figure 0.57. Specimen 06-05-A276c, cervical cancer tissue**

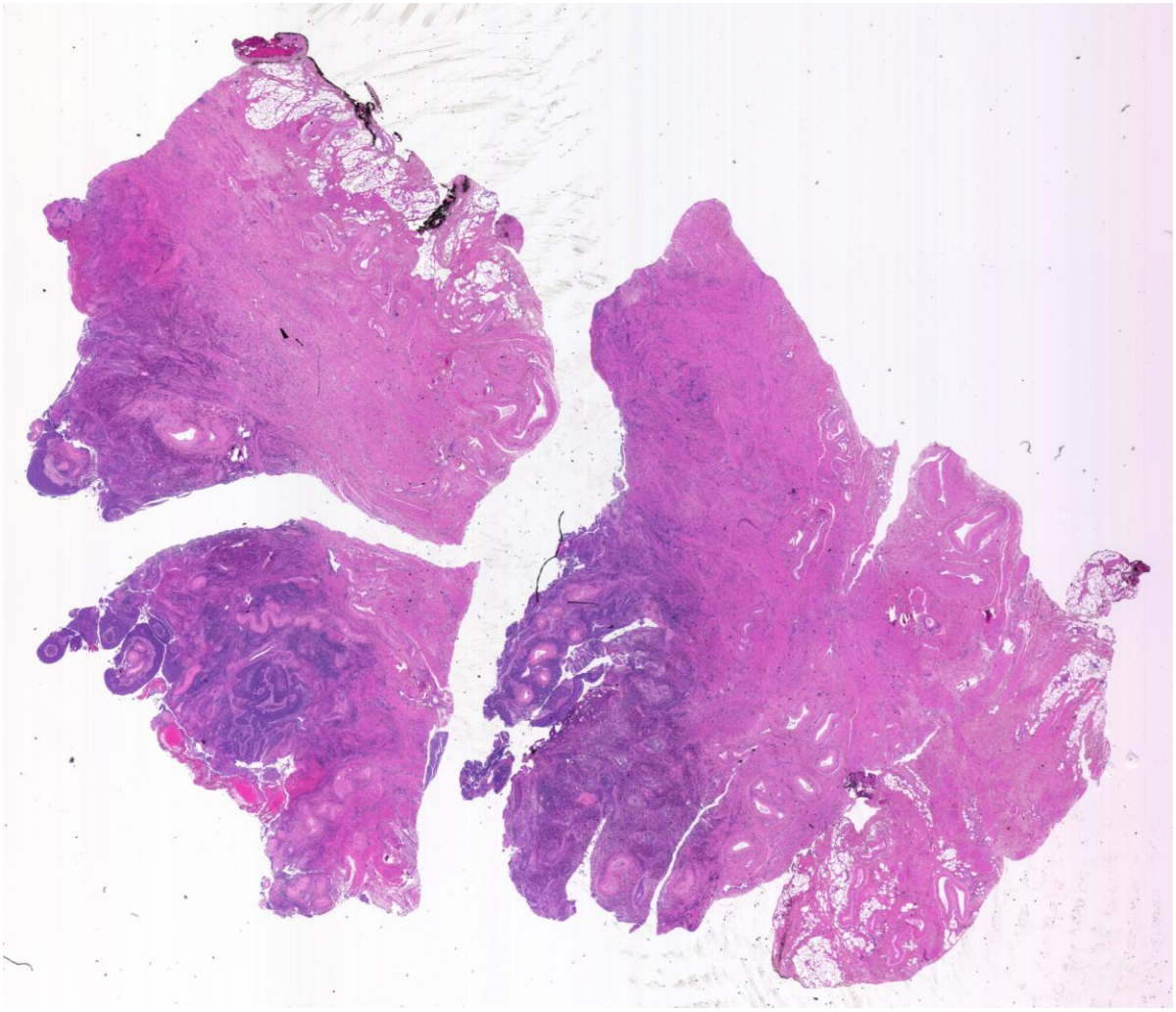


**Figure 0.58. Specimen 06-05-A276c, H&E stain, squamous cell carcinoma of the cervix**

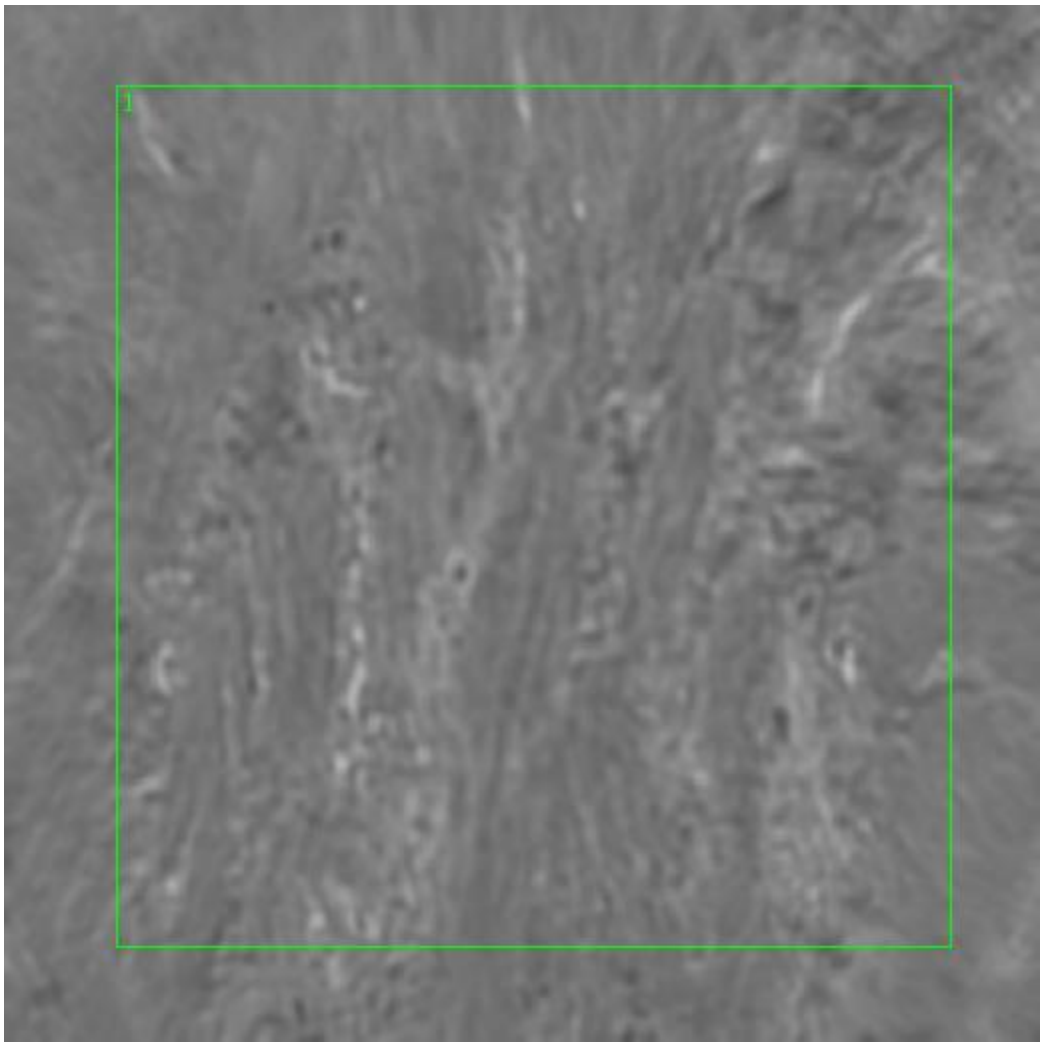



**Figure 0.59. Point A Regions of Interest**
**Region 1: lamina propia**

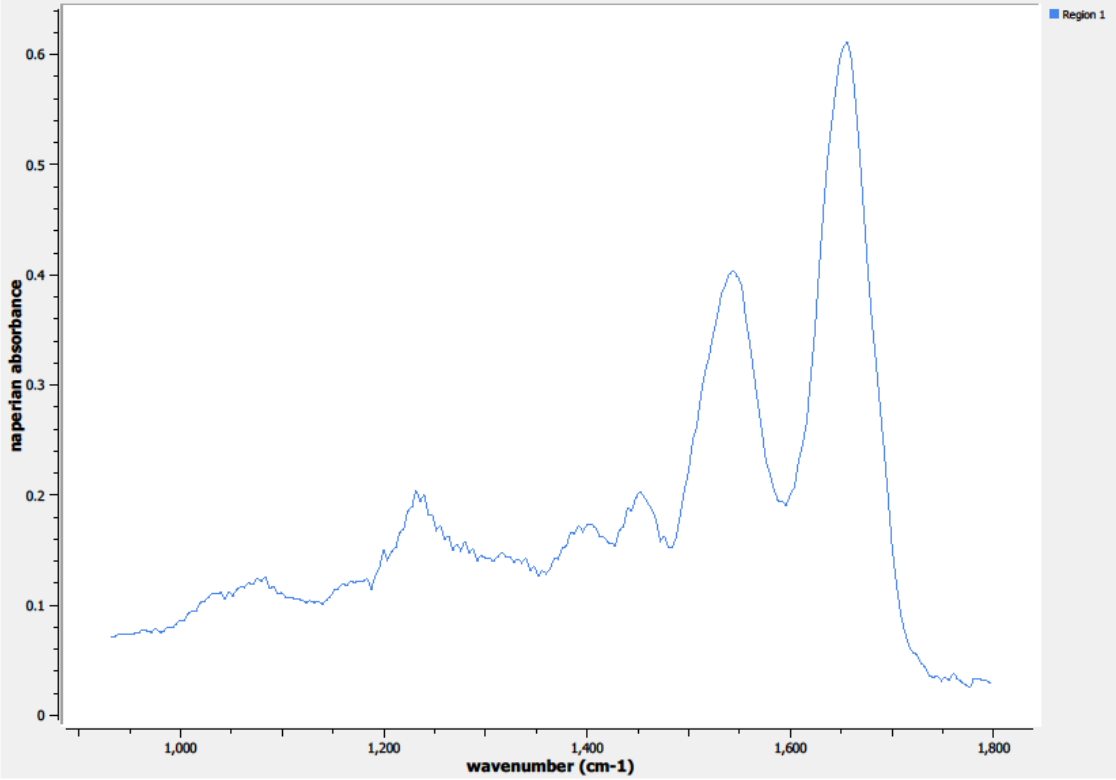



**Figure 0.60. Point A Spectra**

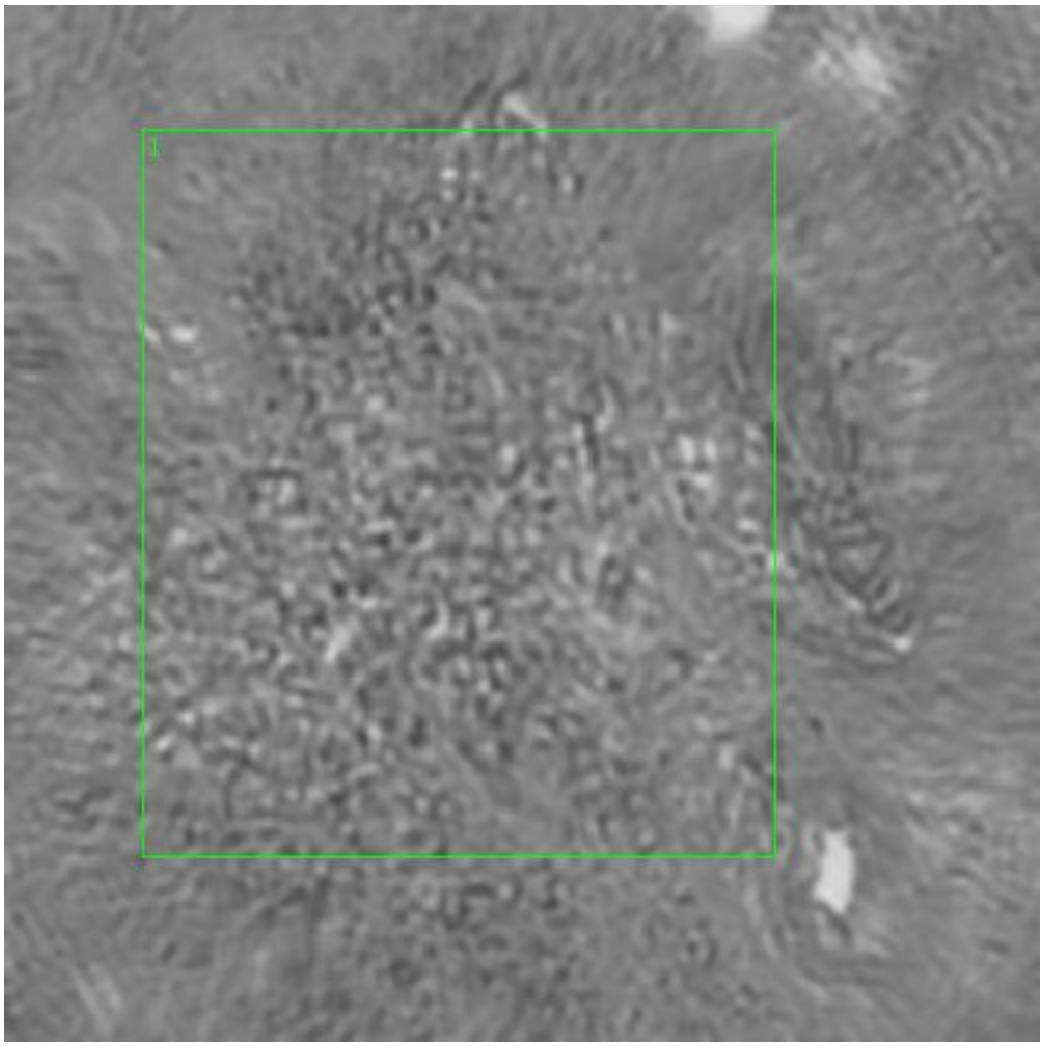


**Figure 0.61. Point V Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

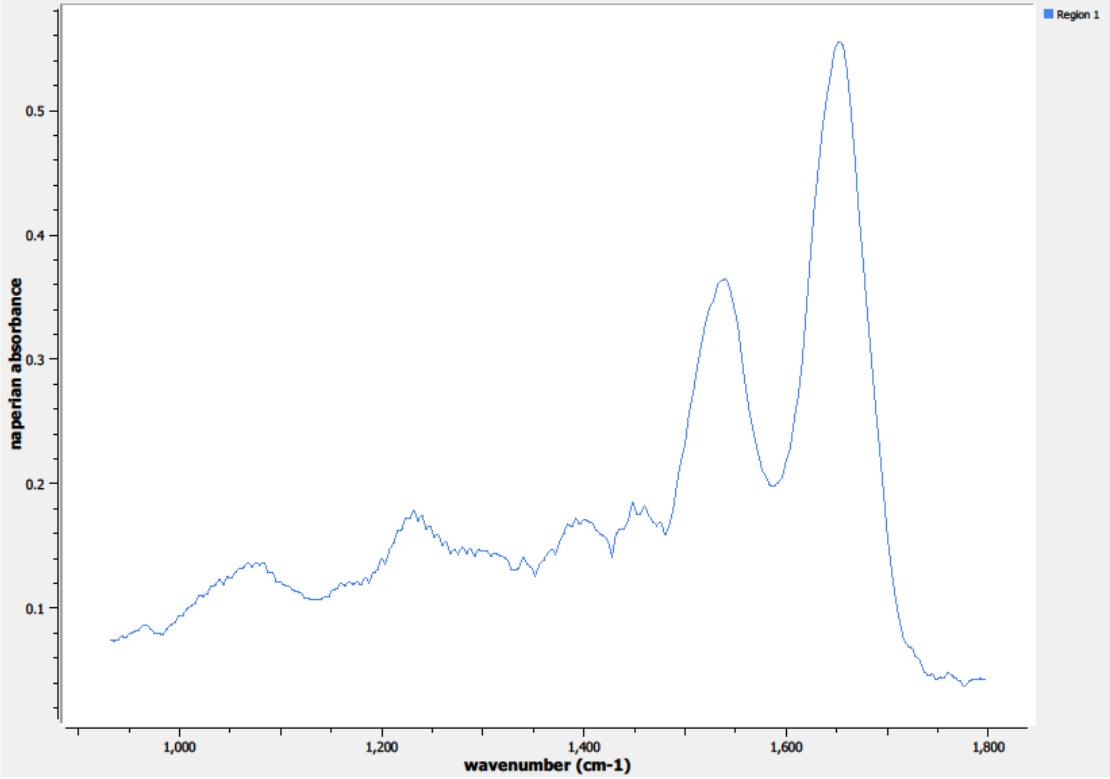



**Figure 0.62. Point V Spectra**

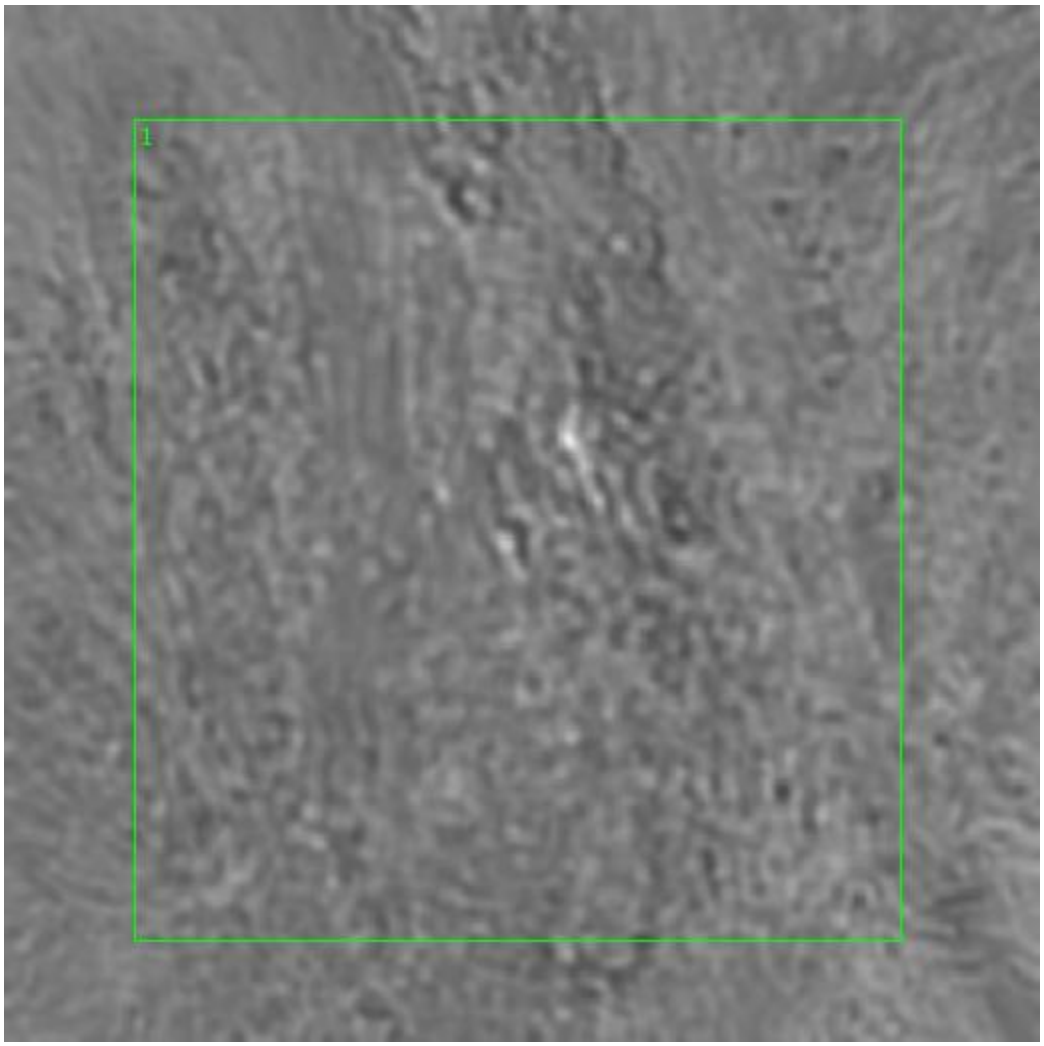



**Figure 0.63. Point X Regions of Interest**
**Region 1: lamina propia, transition region between squamous cell carcinoma and normal tissue**

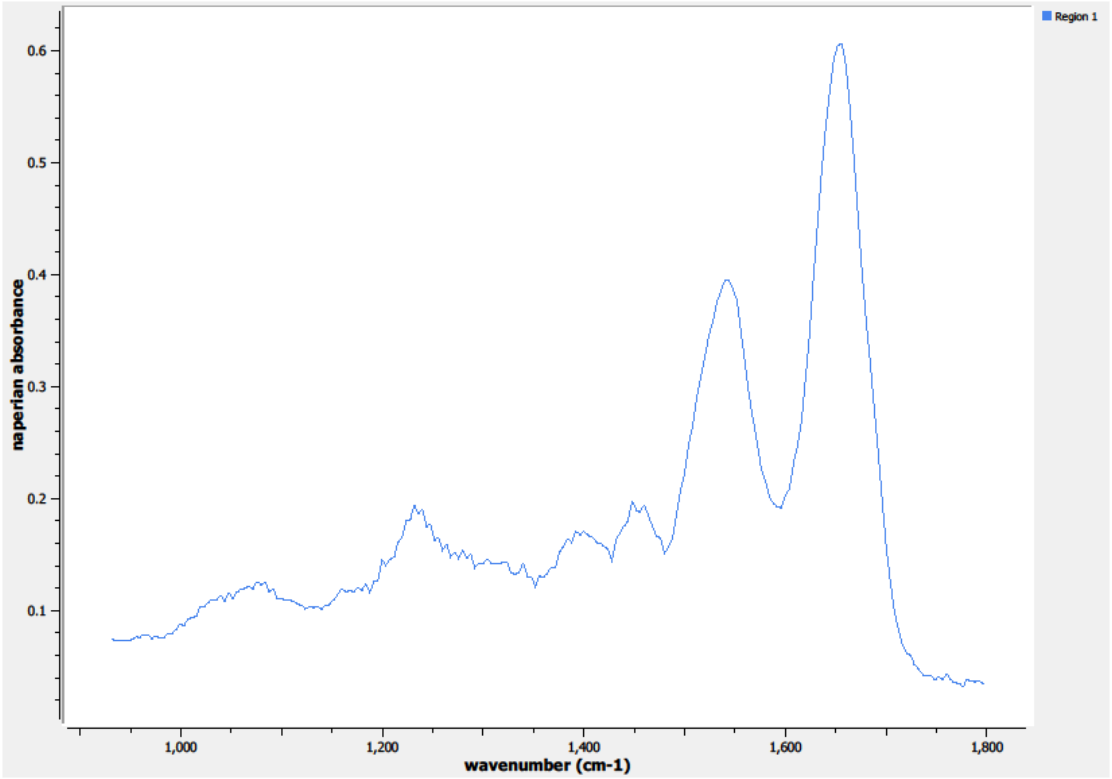



**Figure 0.64. Point X Spectra**

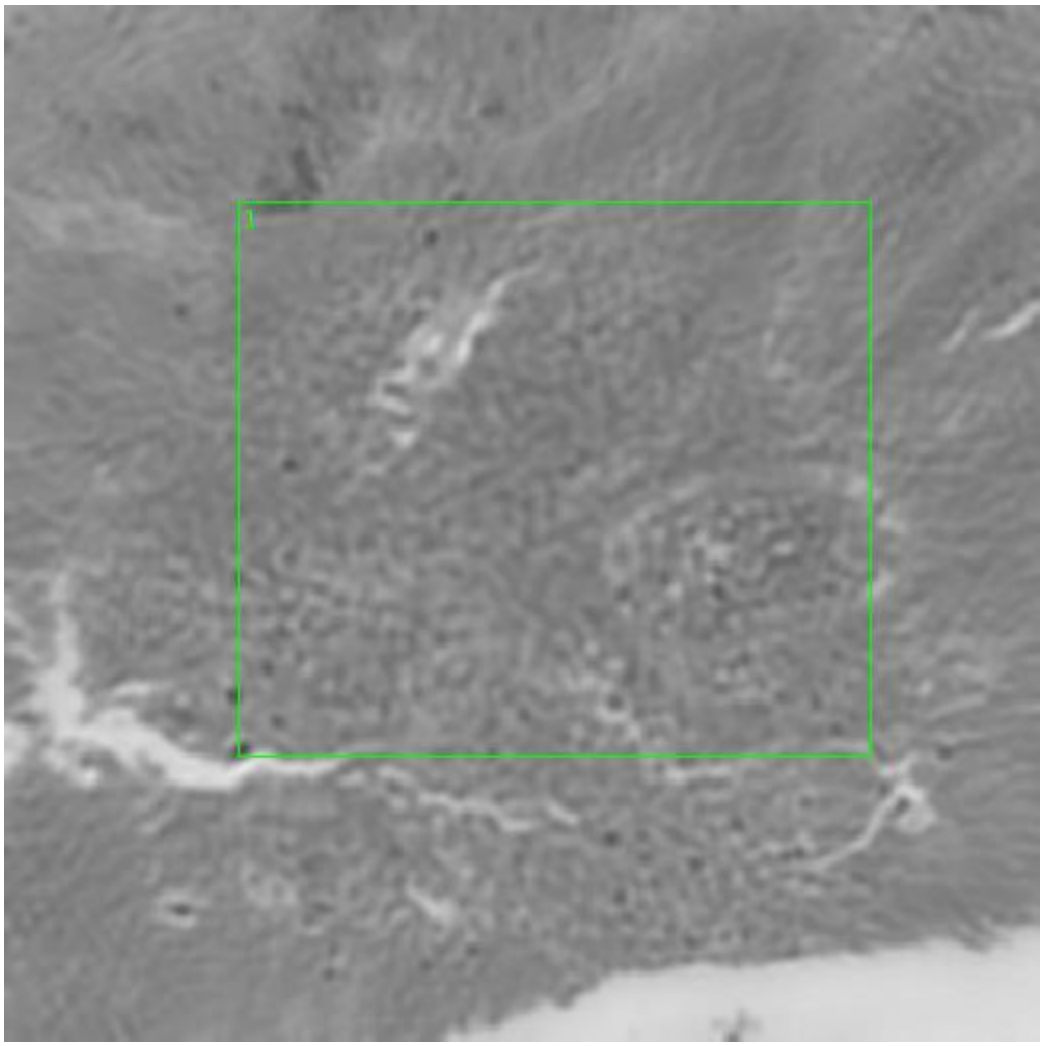



**Figure 0.65. Point Z Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

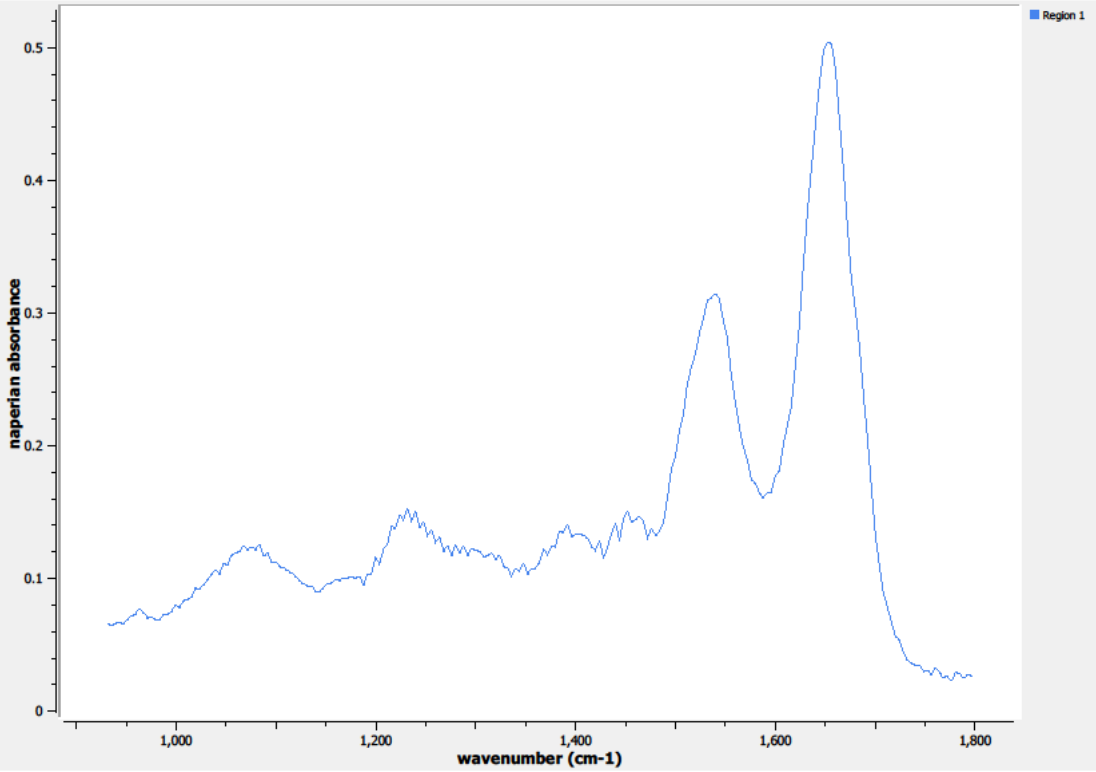



**Figure 0.66. Point Z Spectra**

## A8.06-07-A082b

| | |
|---|---|
| Unique ID | 06-07-A082b |
| Age | 37 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix malignant, adenocarcinoma, primary |

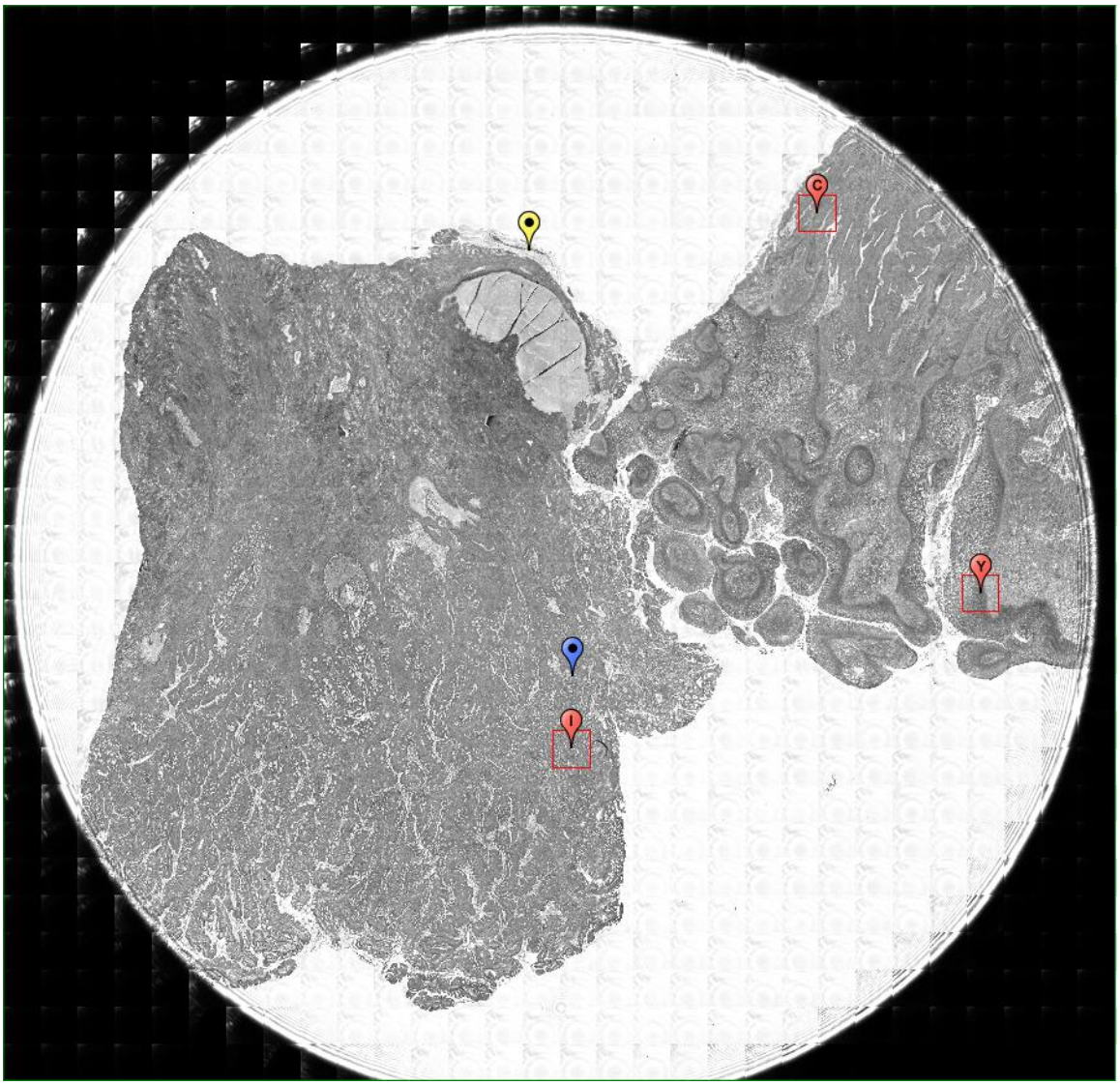



**Figure 0.67. Specimen 06-07-A082b, cervical cancer tissue**

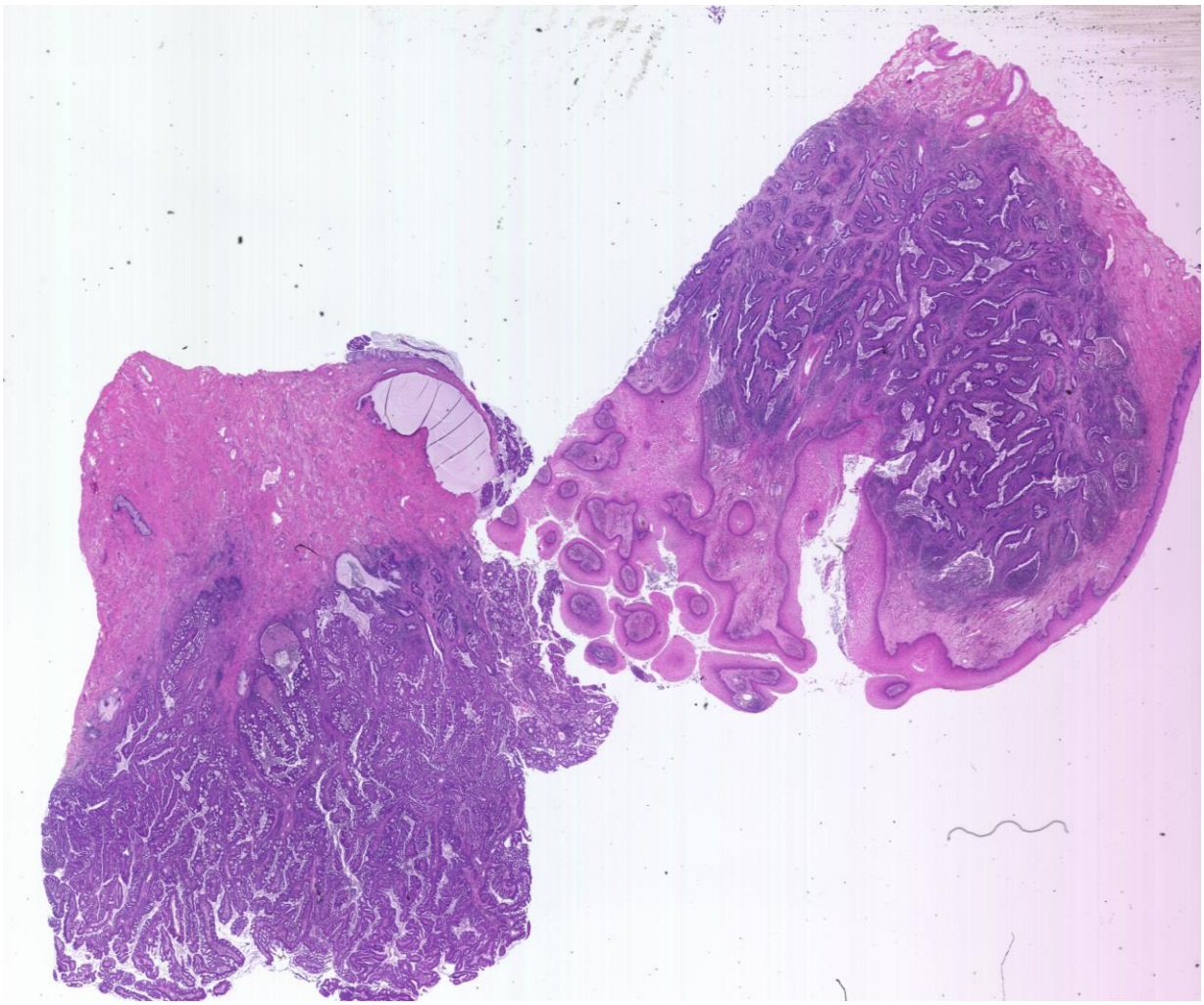


**Figure 0.68. Specimen 06-07-A082b, H&E stain, adenocarcinoma of the cervix**

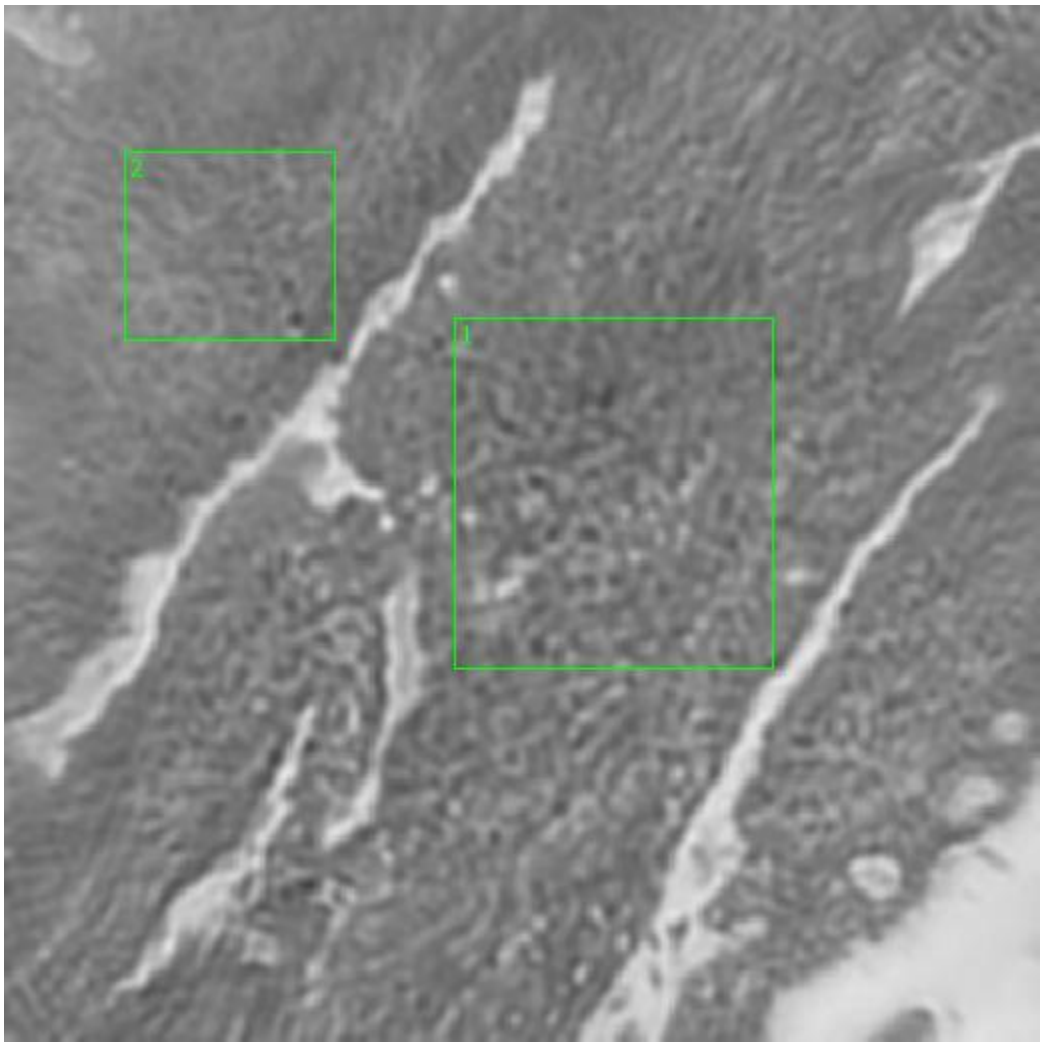



**Figure 0.69. Point C Regions of Interest**
**Regions 1, 2: lamina propia, adenocarcinoma**

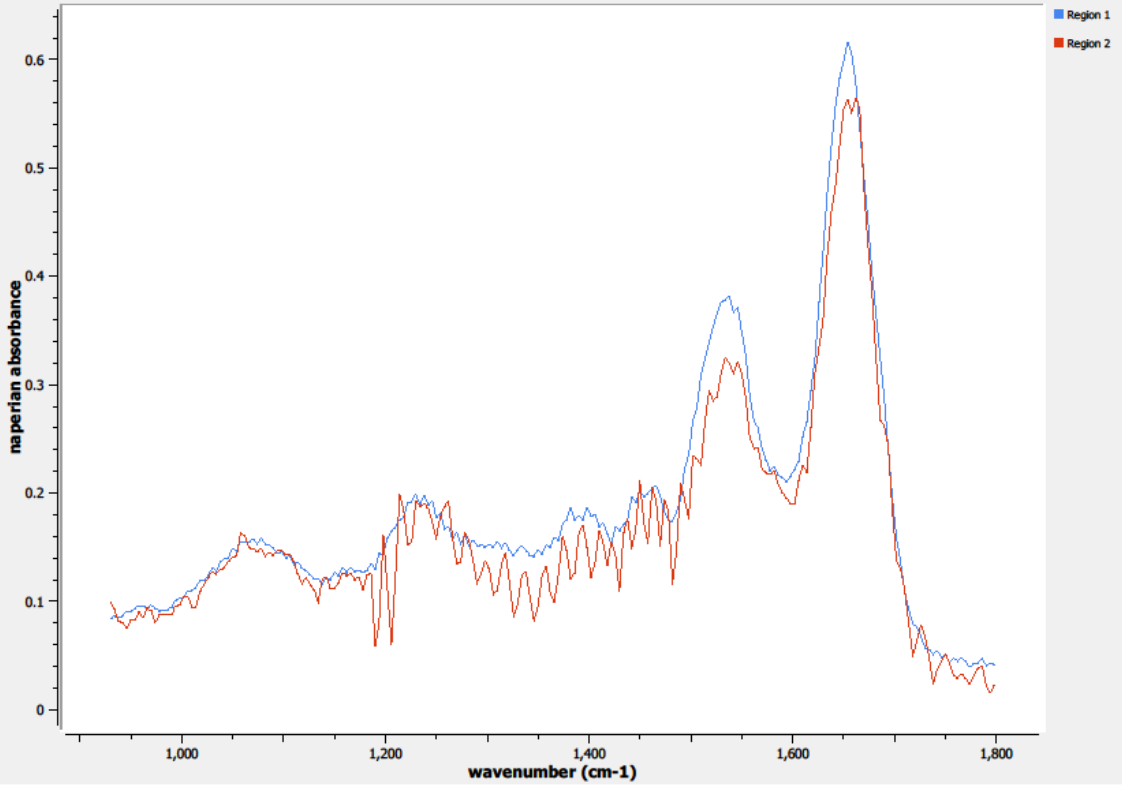



**Figure 0.70. Point C Spectra**

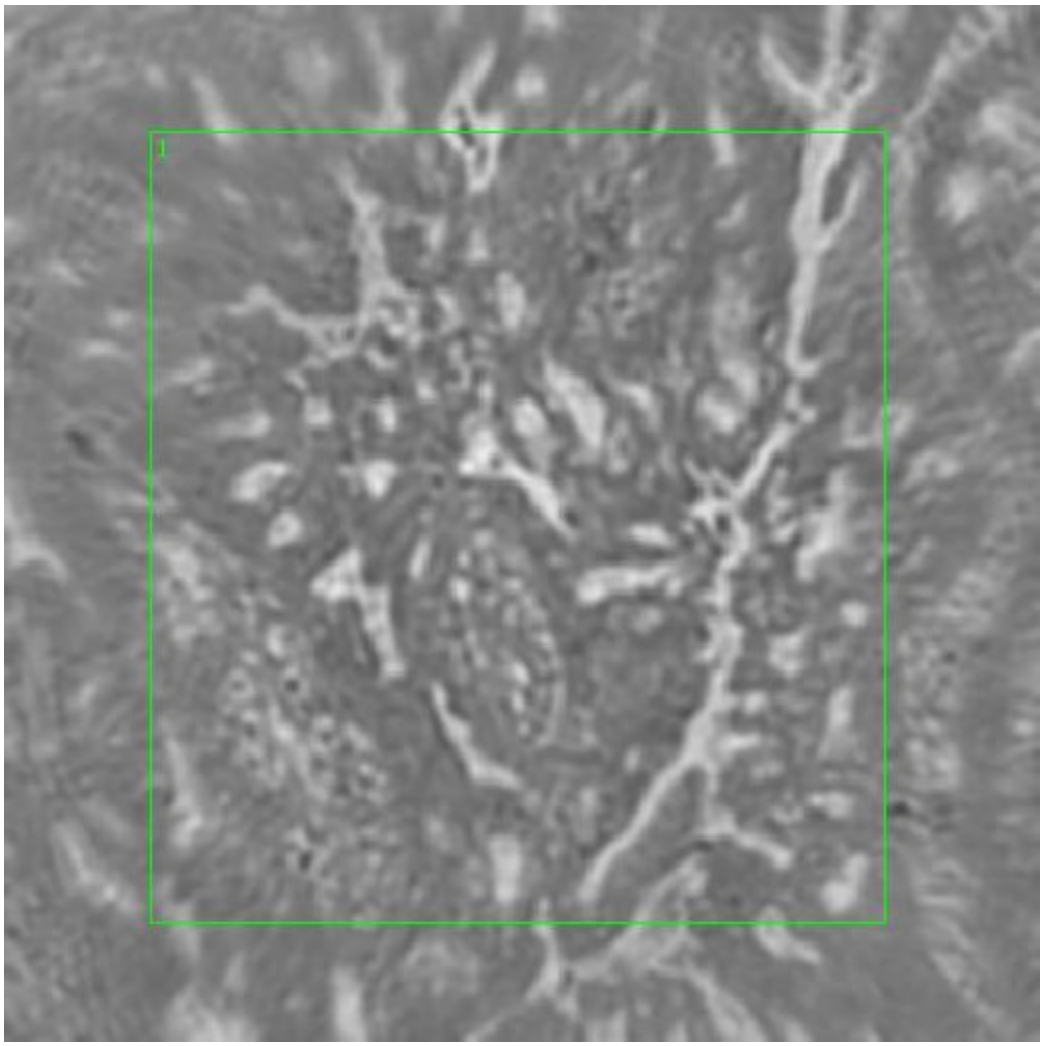



**Figure 0.71. Point I Regions of Interest**
**Region 1: lamina propia, adenocarcinoma**

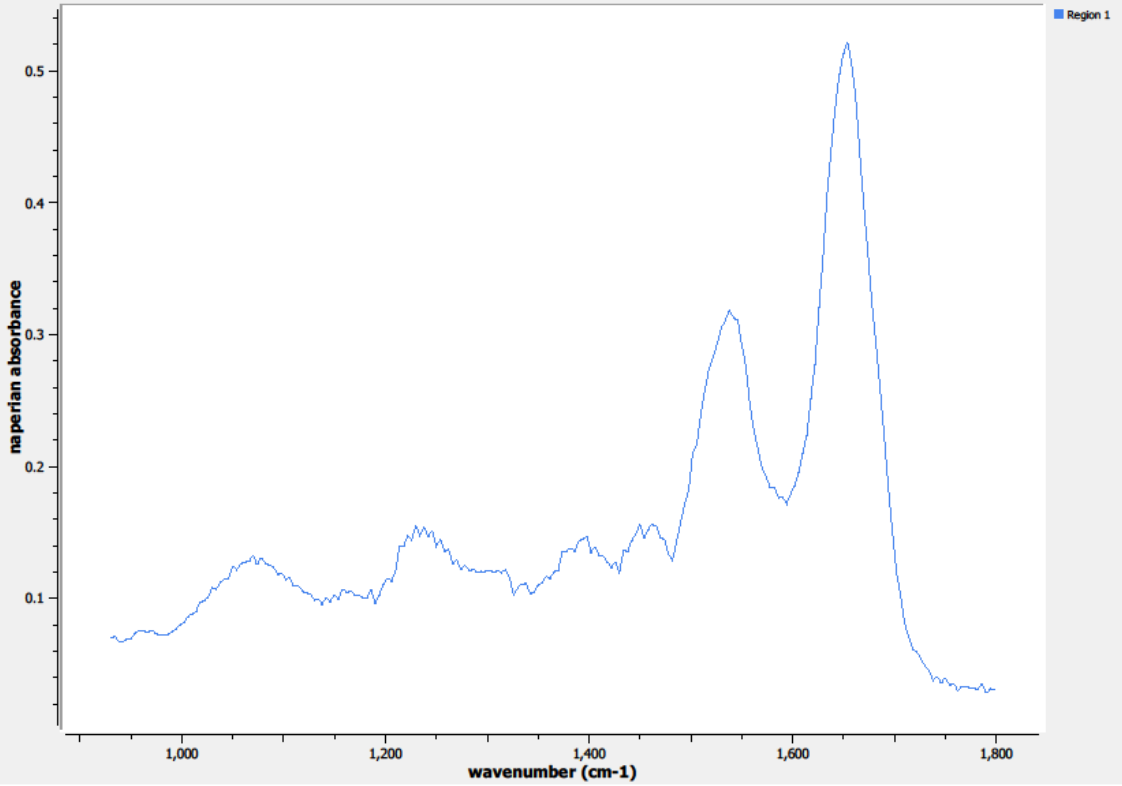



**Figure 0.72. Point I Spectra**

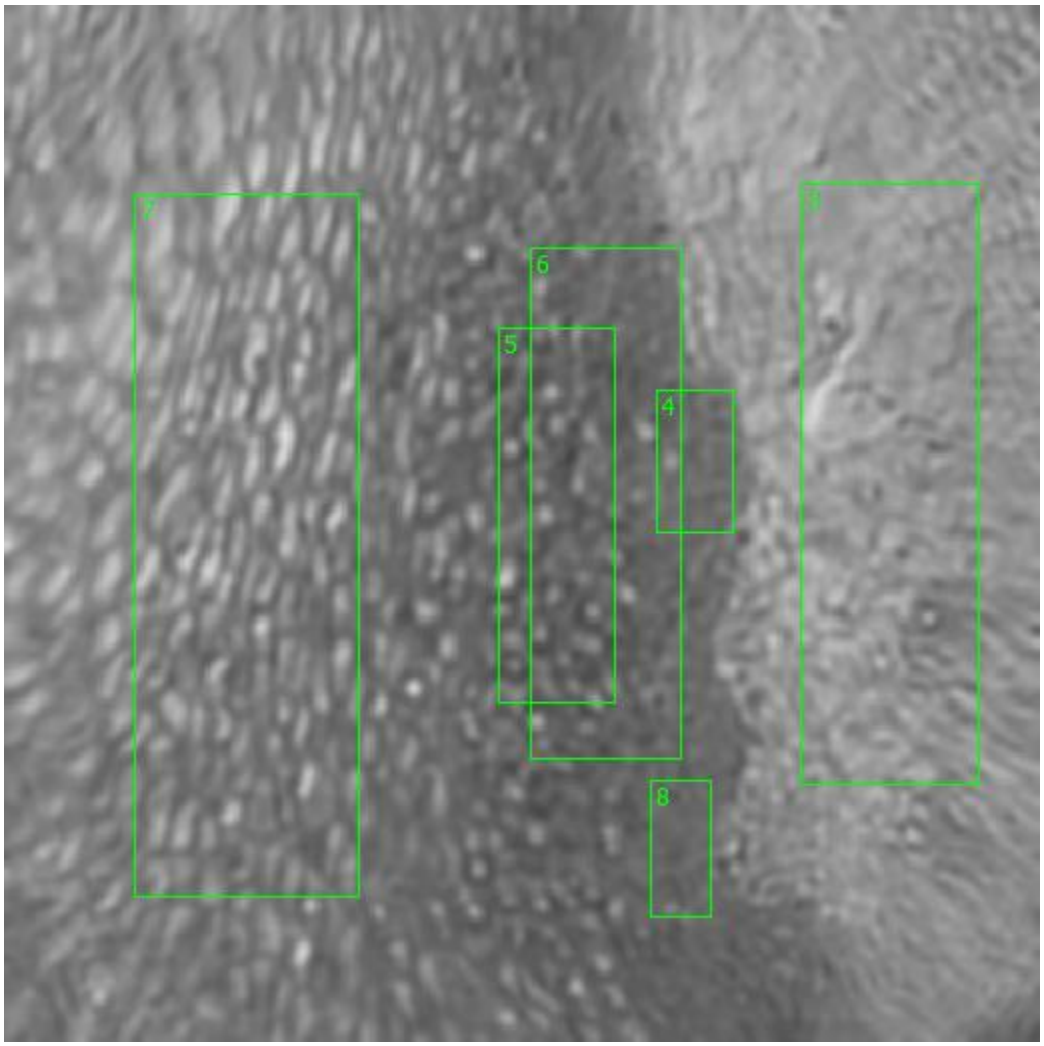



**Figure 0.73. Point Y Regions of Interest**
**Region 3: basement membrane, lamina propia**
**Regions 4, 8: basal layer**
**Region 5: parabasal layer, columnar epithelium**
**Region 6: basal layer, parabasal layer**
**Region 7: stratified squamous epithelium**

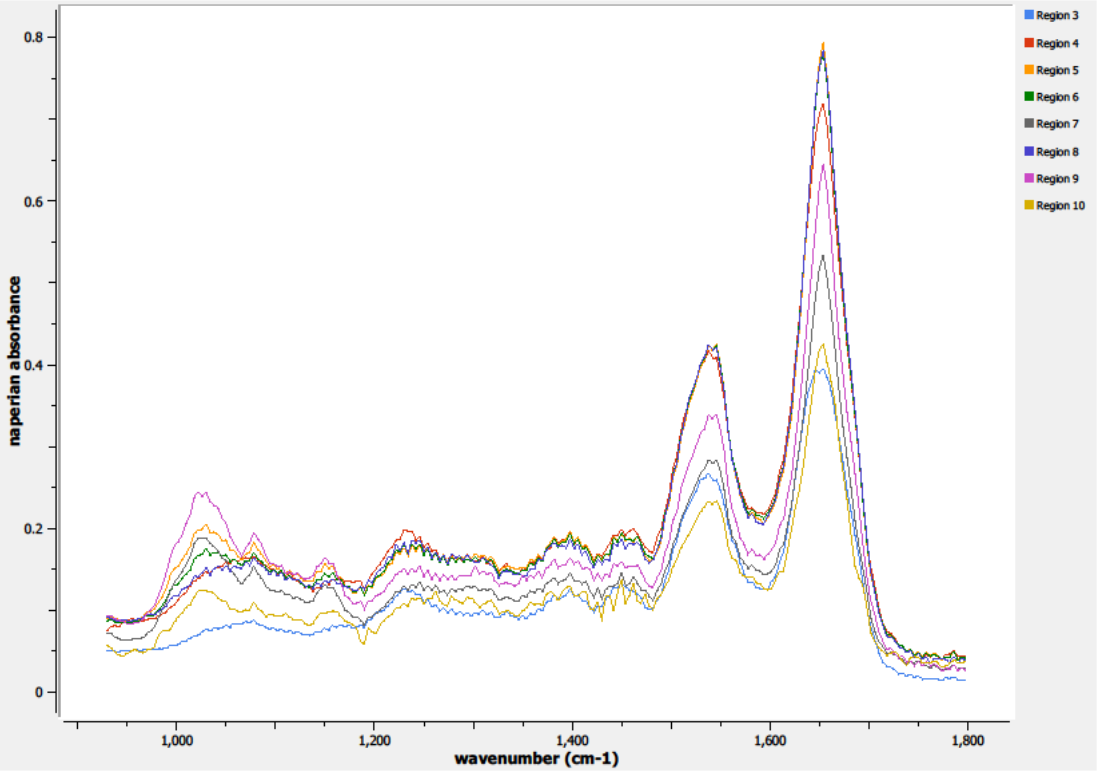



**Figure 0.74. Point Y Spectra**

## A9.10-09-A121a

| | |
|---|---|
| Unique ID | 10-09-A121a |
| Age | 30 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | primary malignant carcinoma, glassy cell, %T (90) / %N (10) |

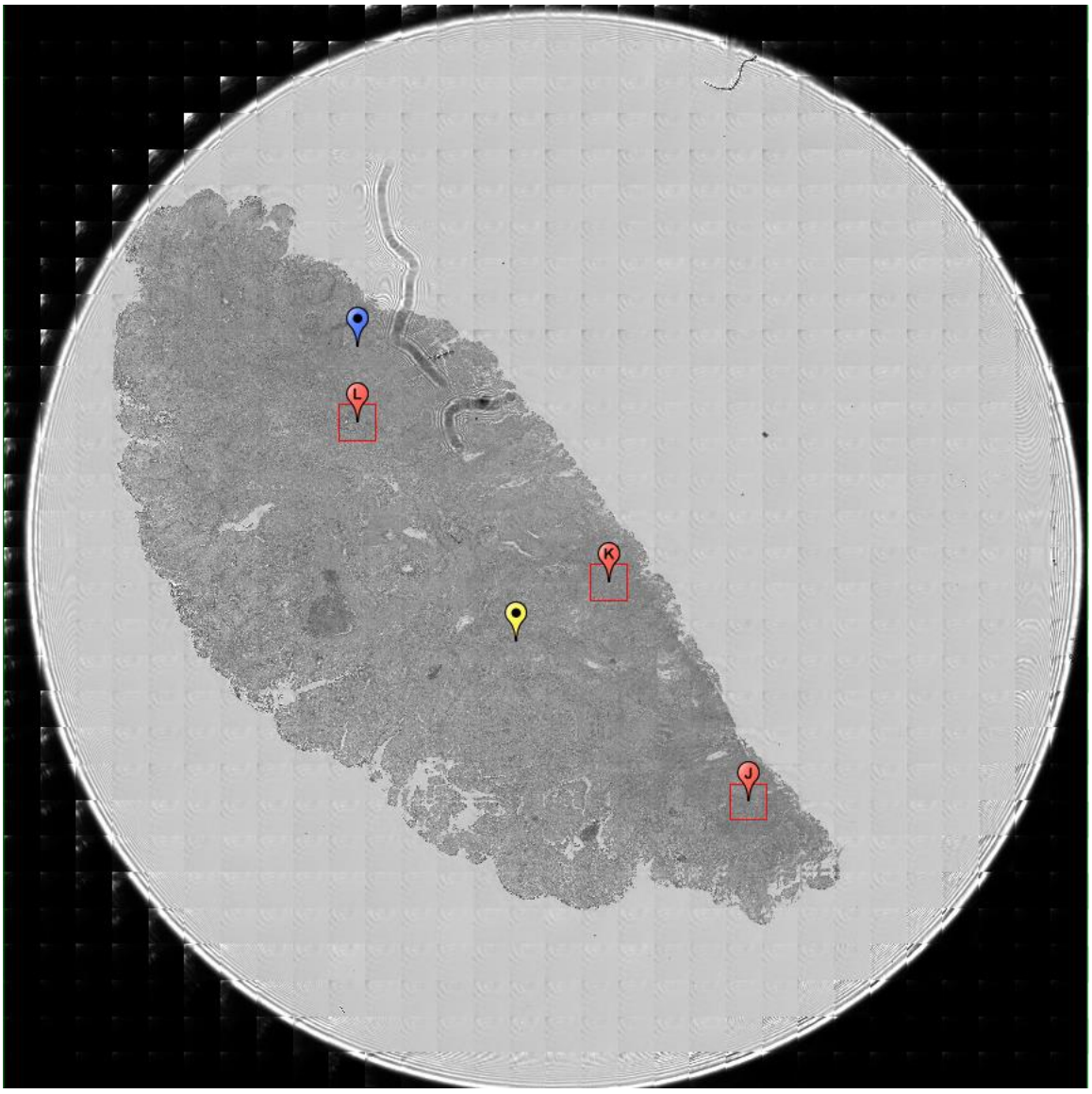



**Figure 0.75. Specimen 10-09-A121a, cervical cancer tissue**

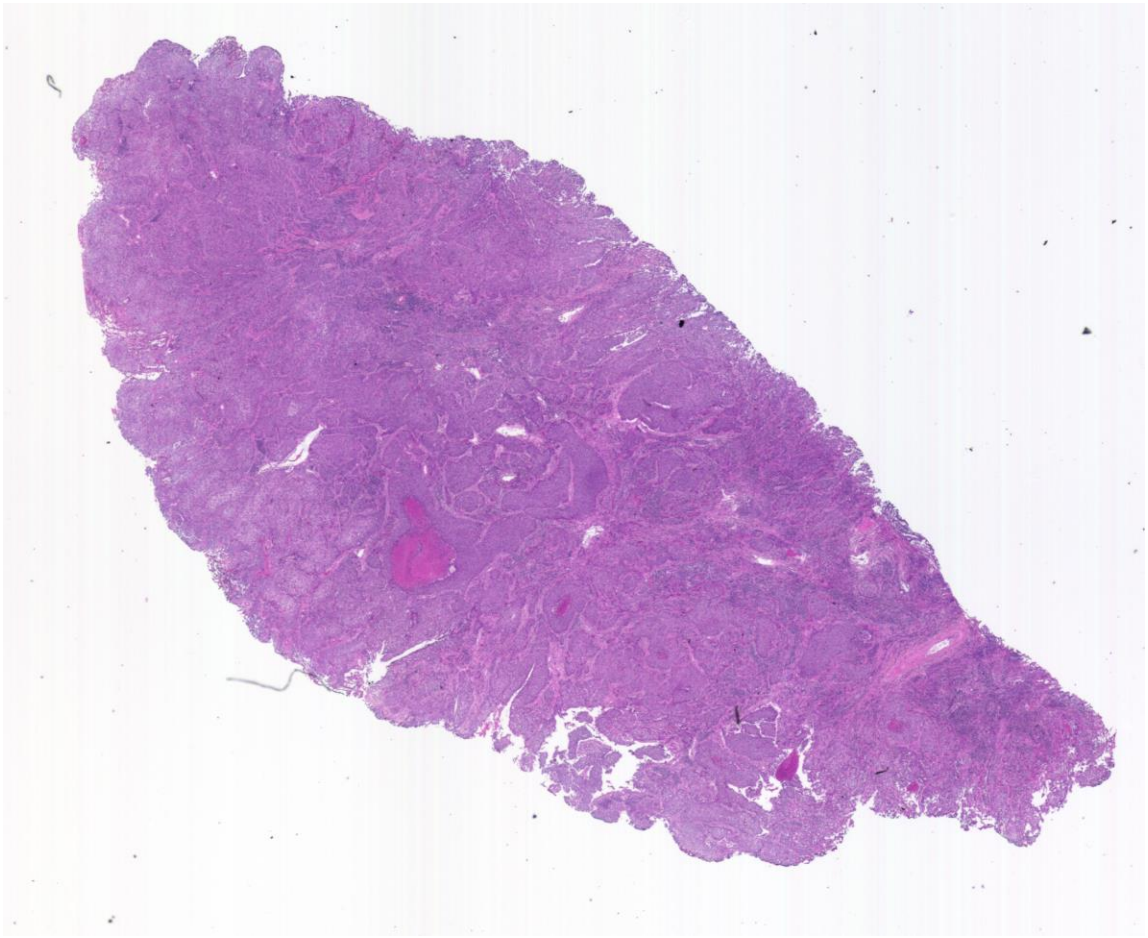


**Figure 0.76. Specimen 10-09-A121a, H&E stain, glassy cell carcinoma of the cervix**

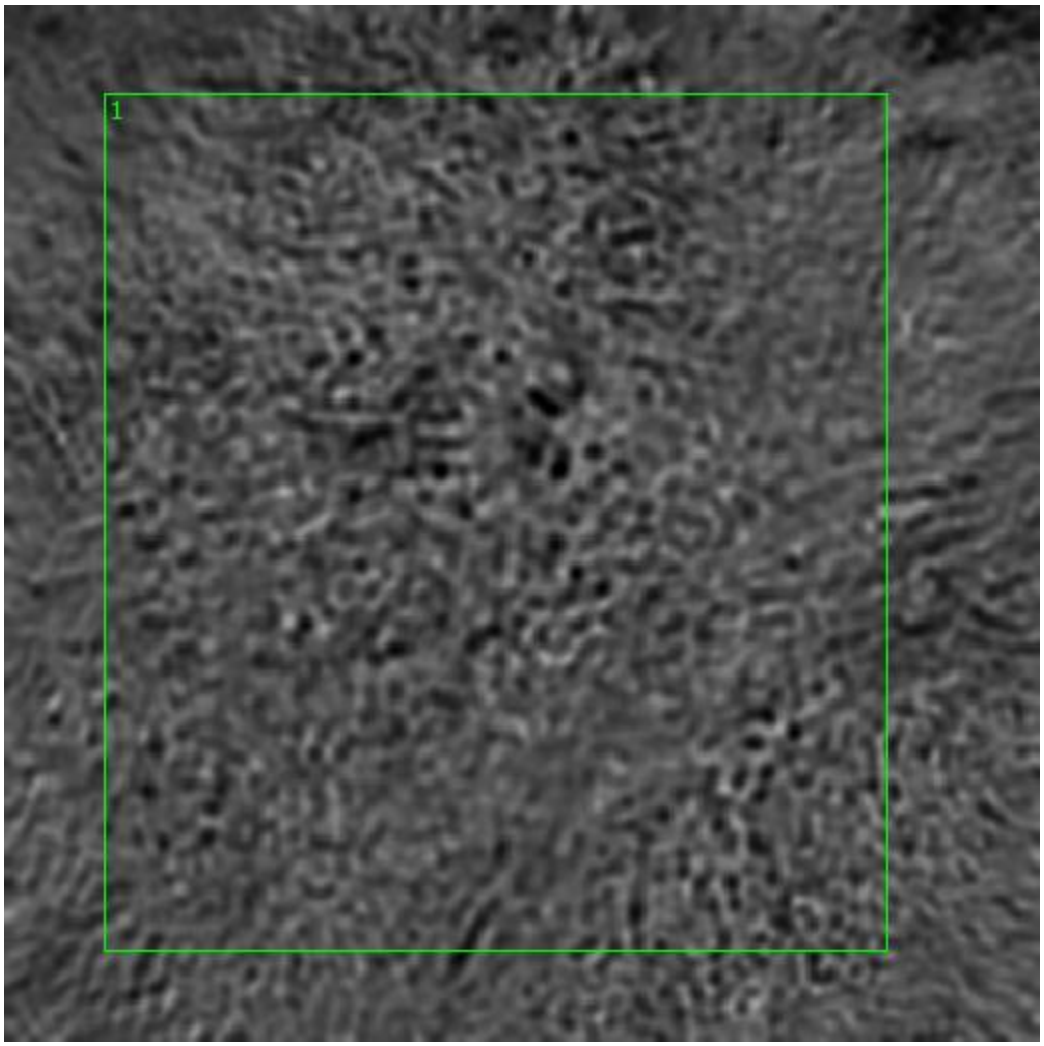



**Figure 0.77. Point J Regions of Interest**
**Region 1: lamina propia, glassy cell carcinoma**

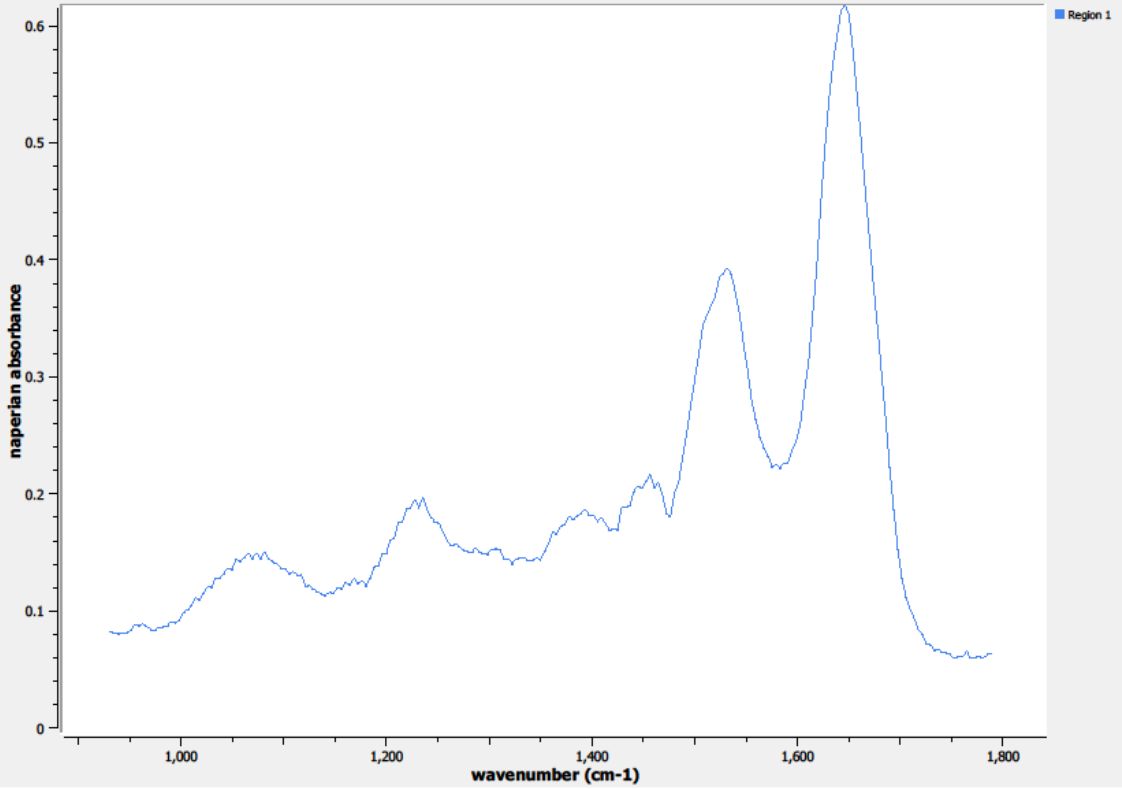



**Figure 0.78. Point J Spectra**

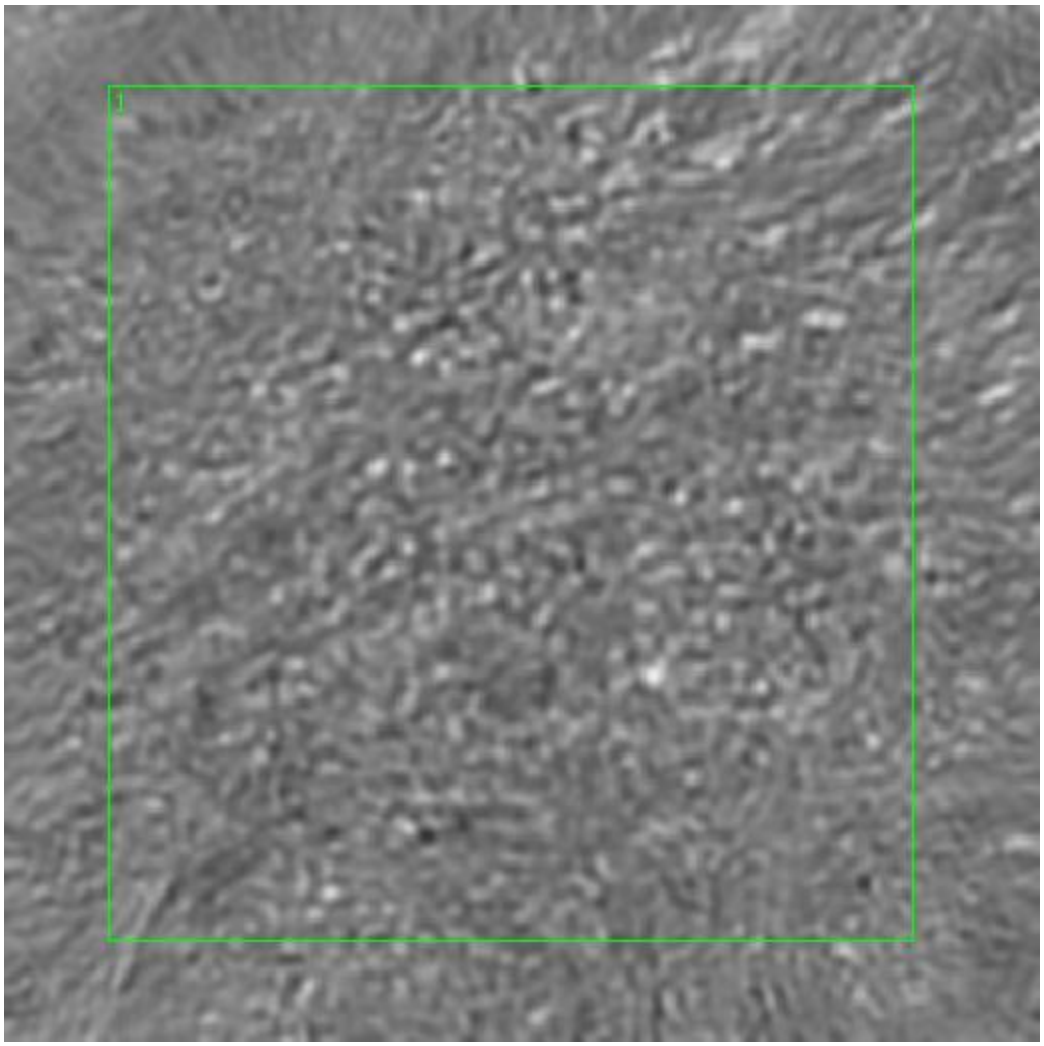



**Figure 0.79. Point K Regions of Interest**
**Region 1: lamina propia, glassy cell carcinoma**

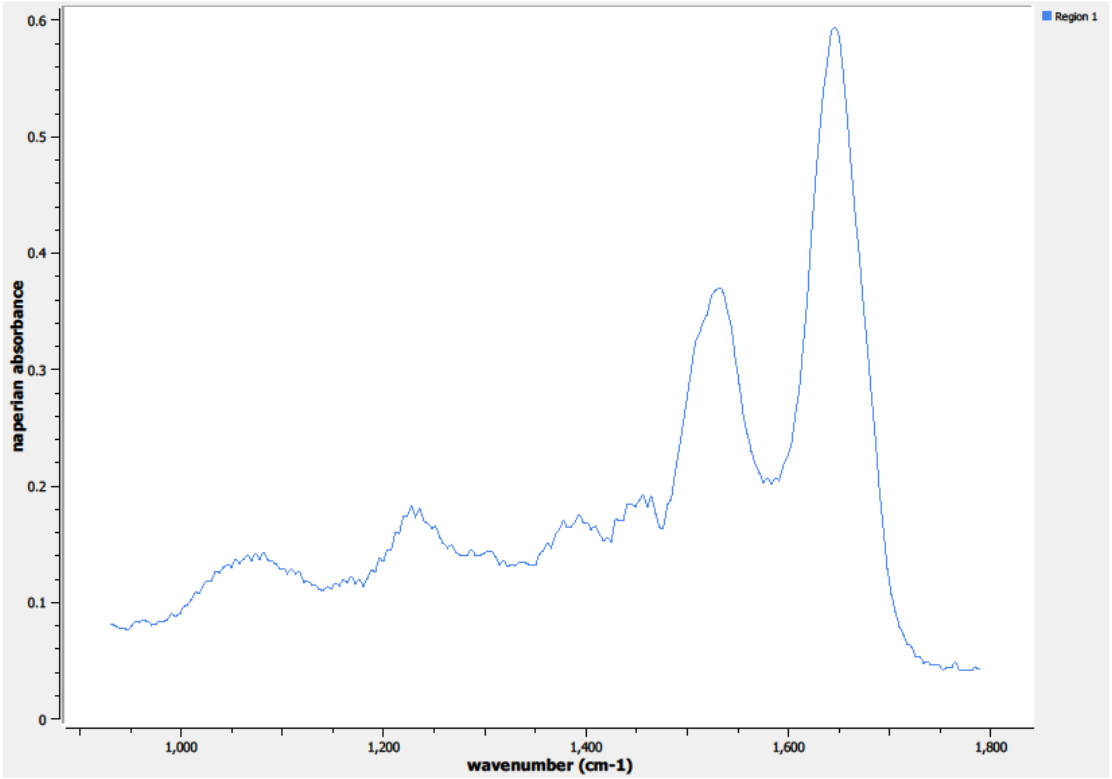



**Figure 0.80. Point K Spectra**

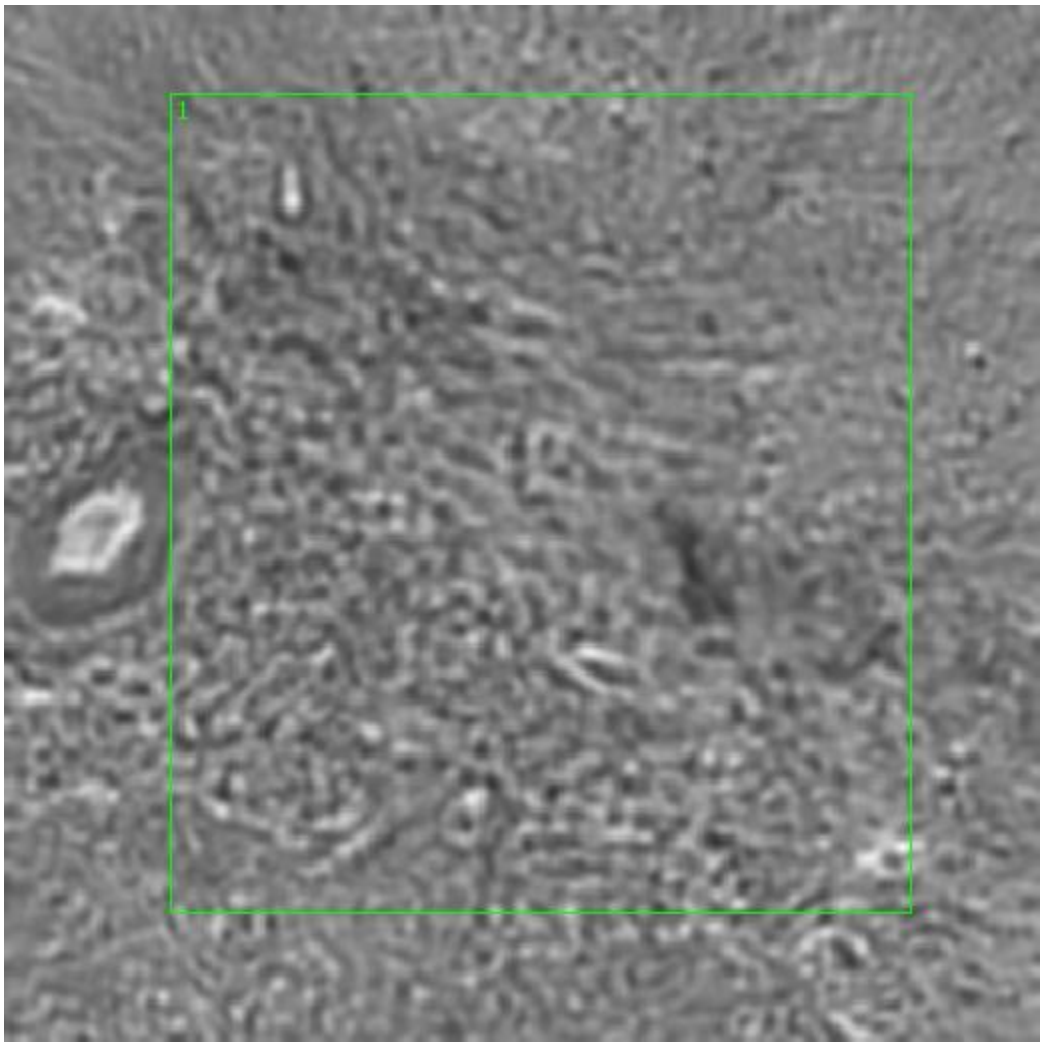



**Figure 0.81. Point L Regions of Interest**
**Region 1: lamina propia, glassy cell carcinoma**

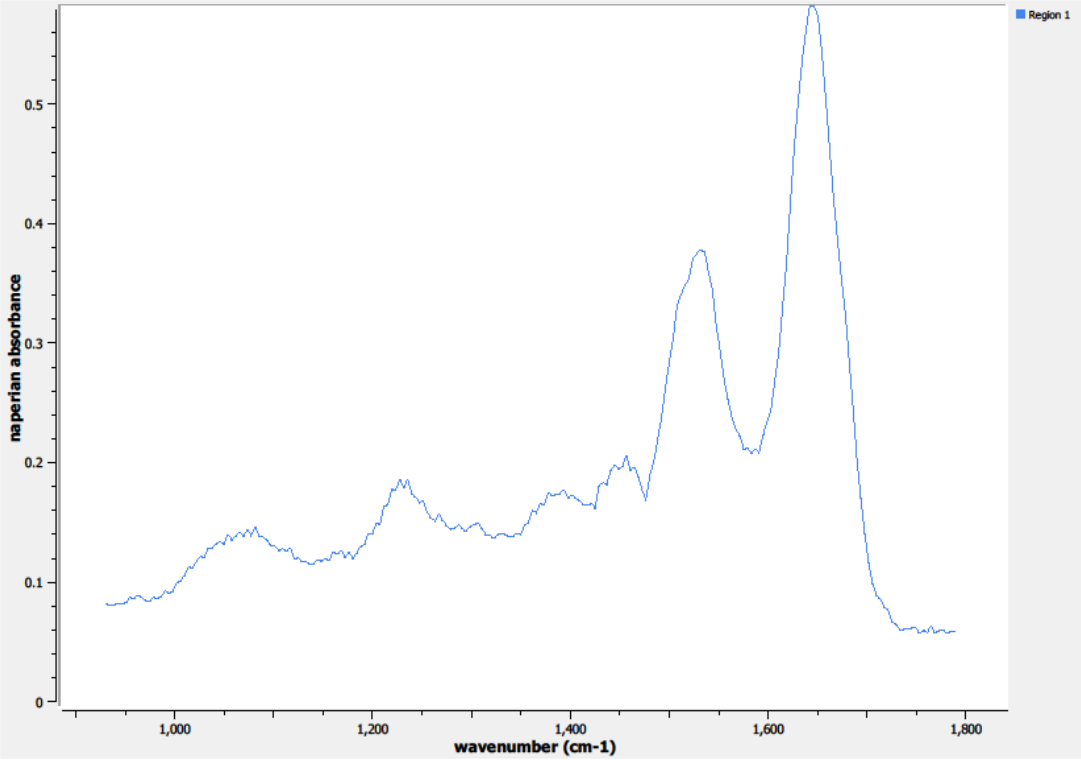



**Figure 0.82. Point L Spectra**

## A10. 10-09-A127b

| | |
|---|---|
| Unique ID | 10-09-A127b |
| Age | 34 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix malignant, squamous cell carcinoma, primary, %T (70) / %N (40) |

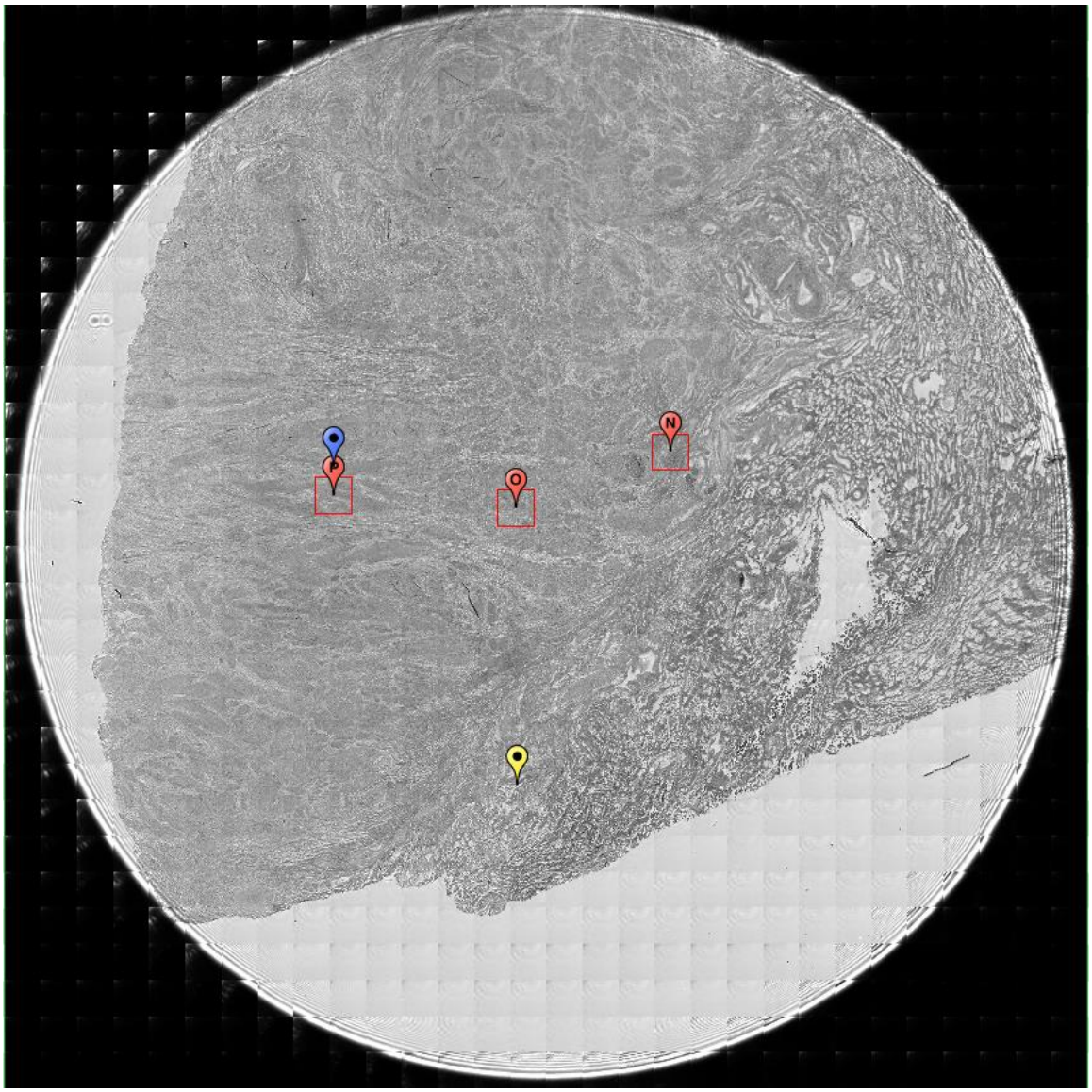



**Figure 0.83. Specimen 10-09-A127b, cervical cancer tissue**

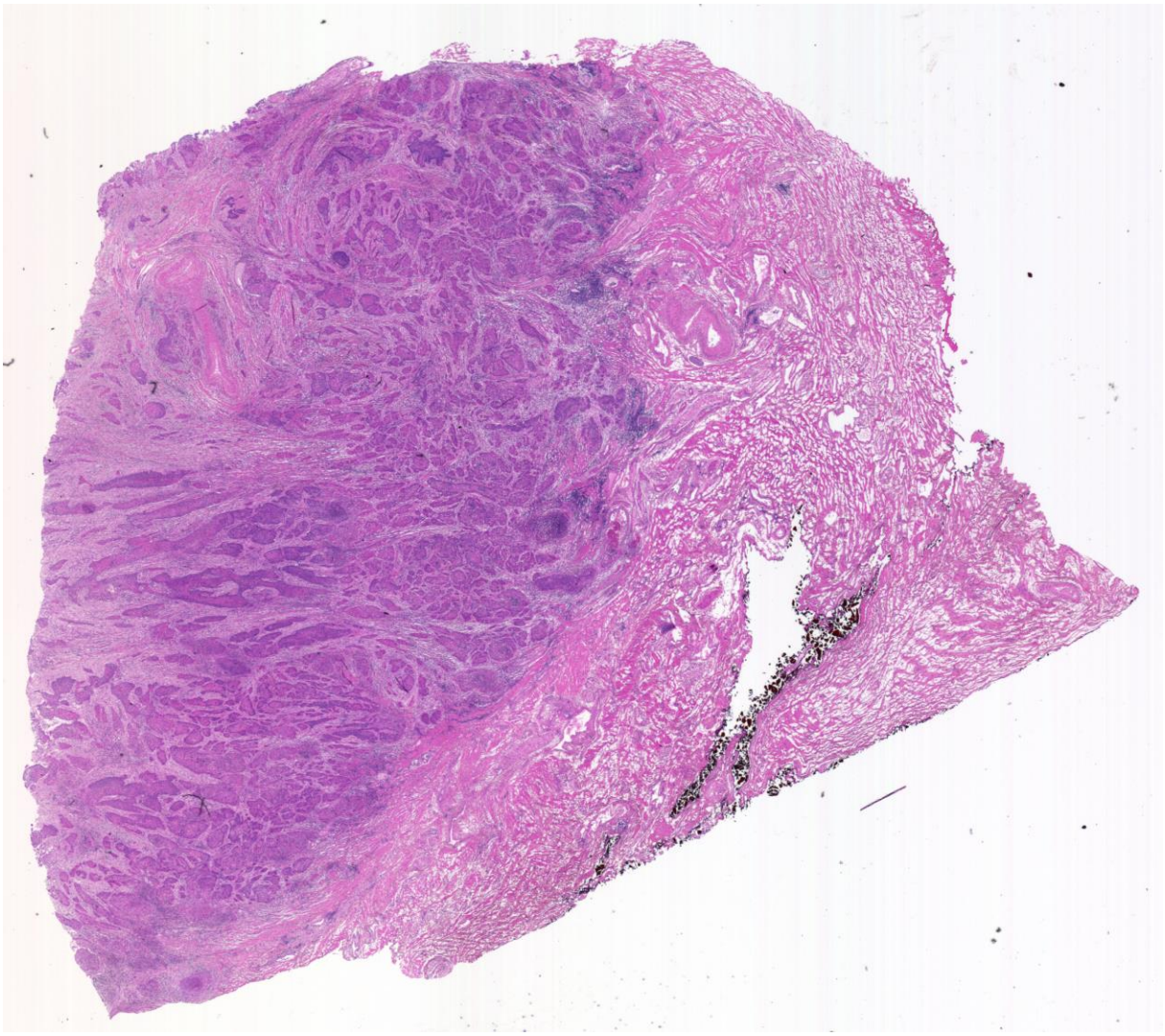


**Figure 0.84. Specimen 10-09-A127b, H&E stain, squamous cell carcinoma of the cervix**

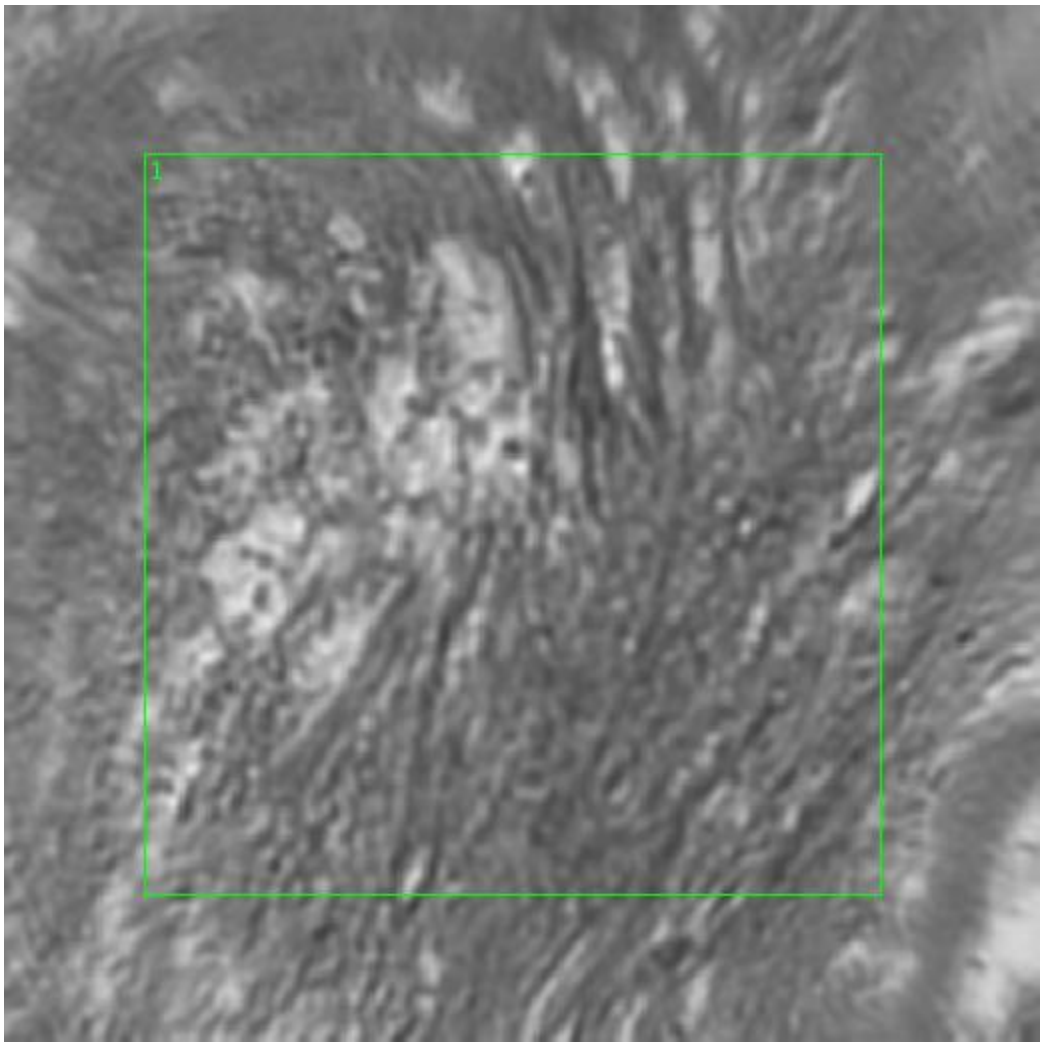



**Figure 0.85. Point N Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

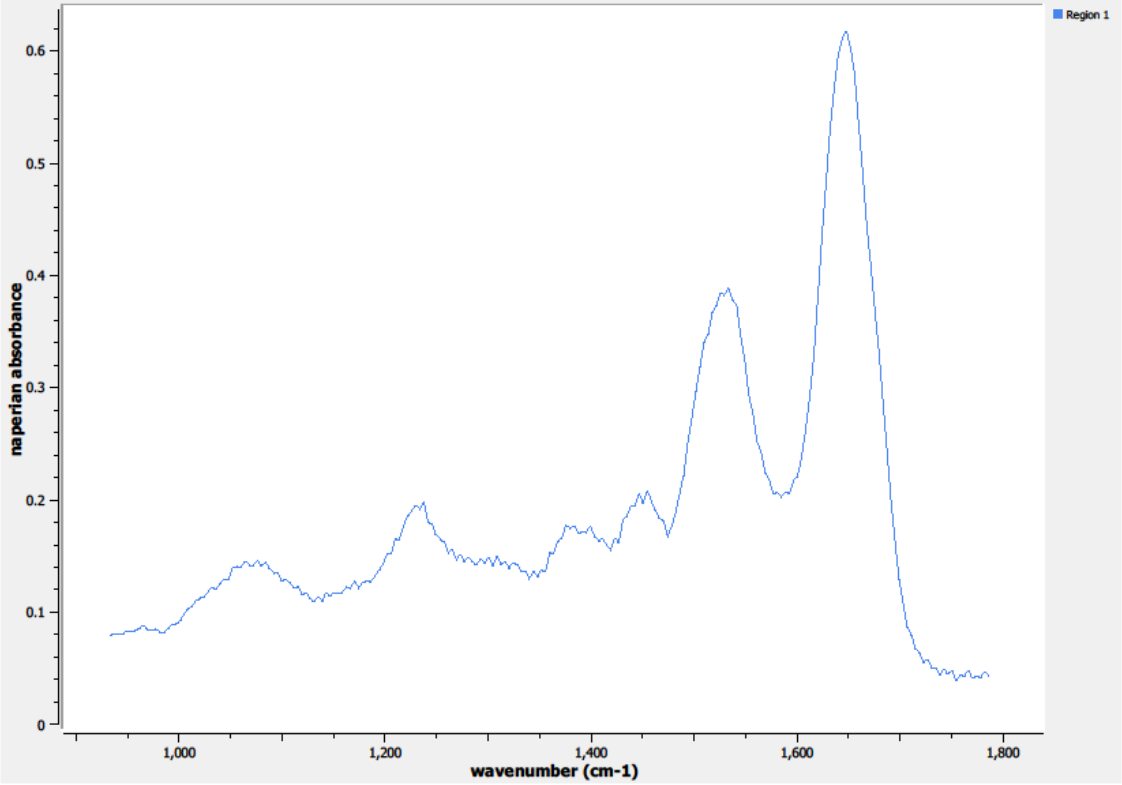



**Figure 0.86. Point N Spectra**

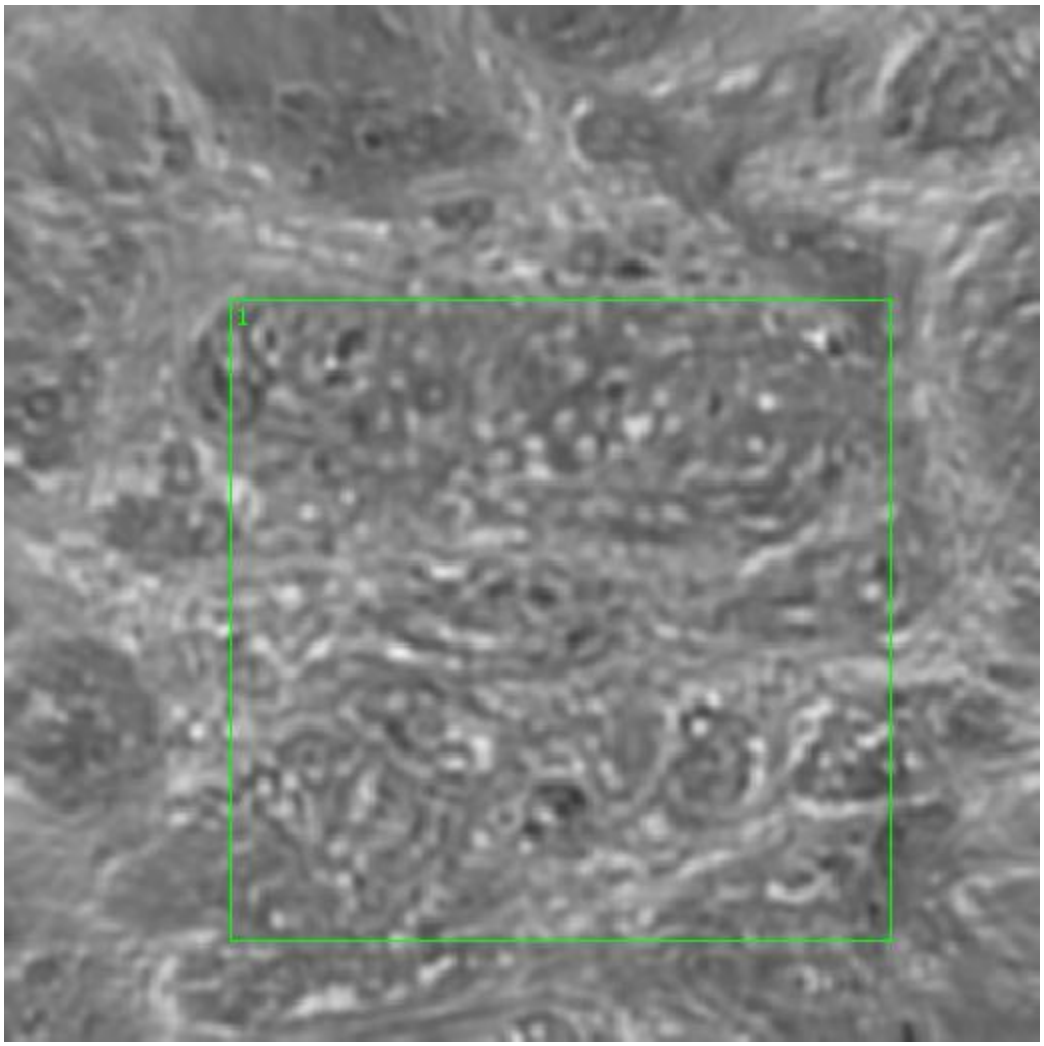



**Figure 0.87. Point O Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

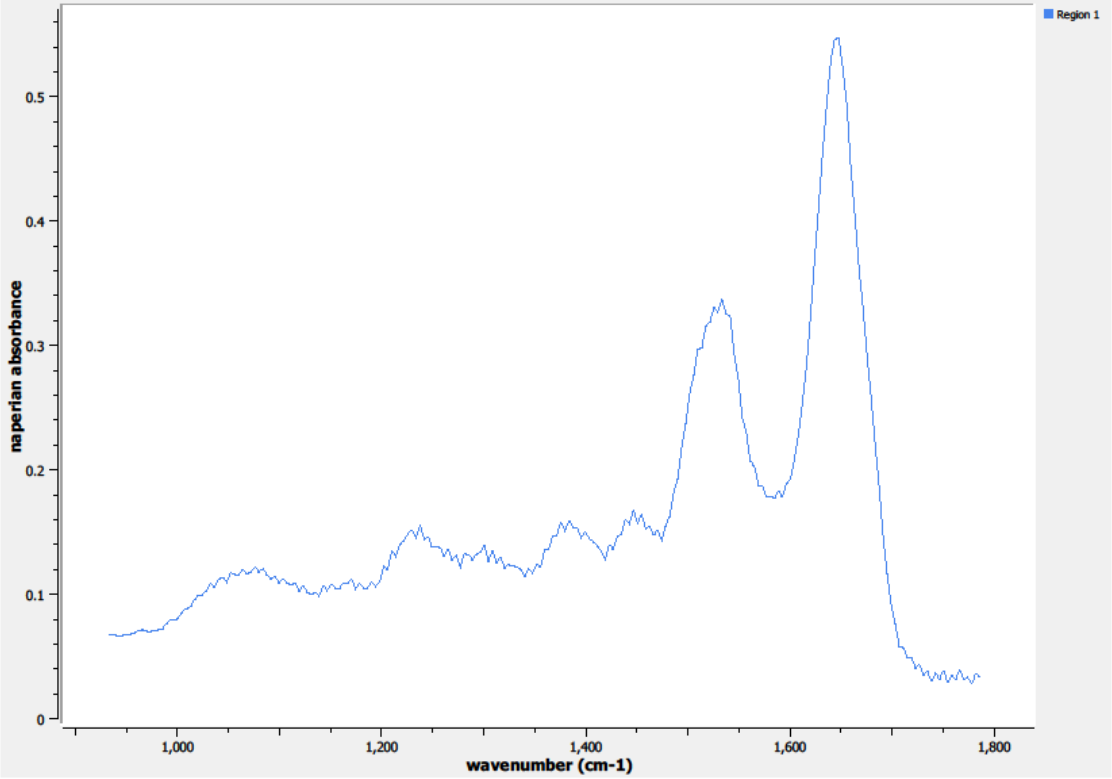



**Figure 0.88. Point O Spectra**

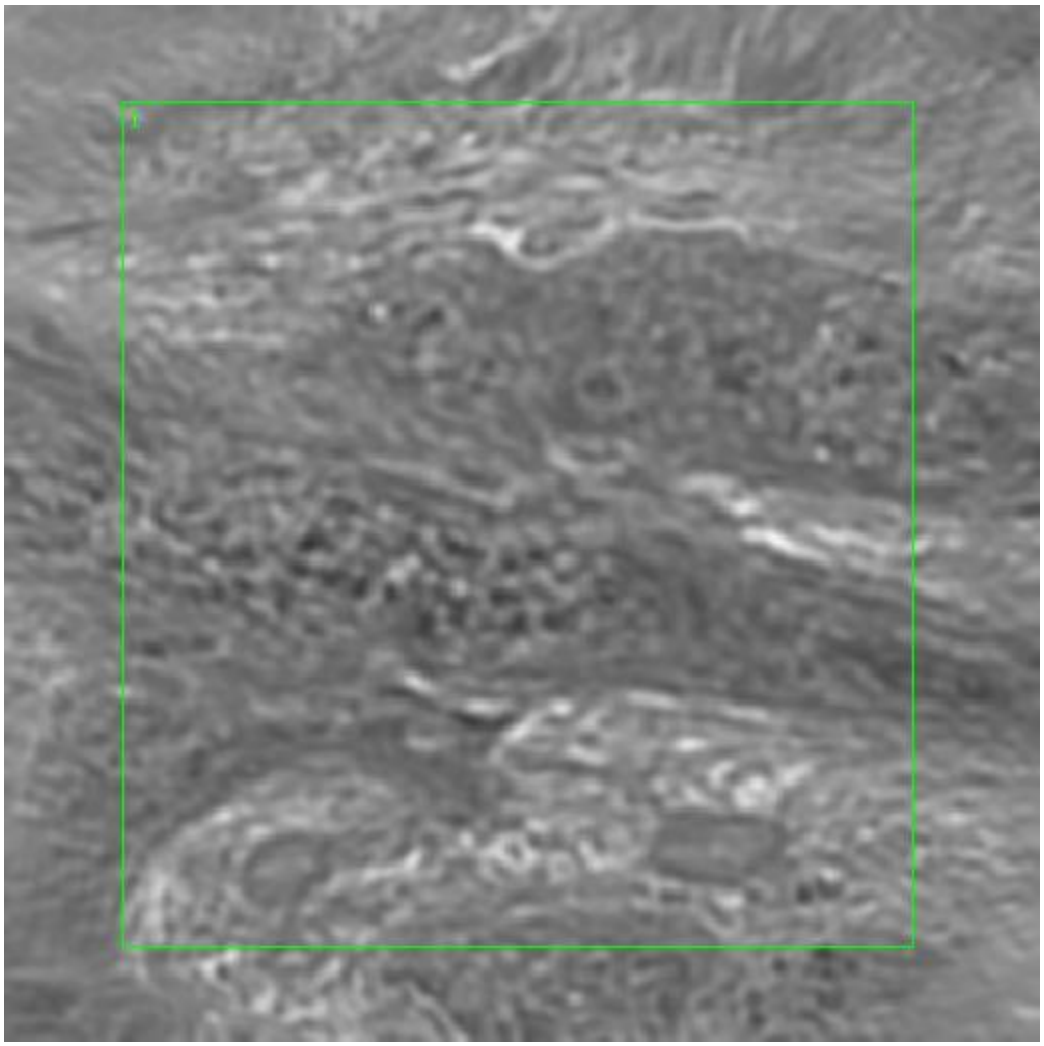


**Figure 0.89. Point P Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

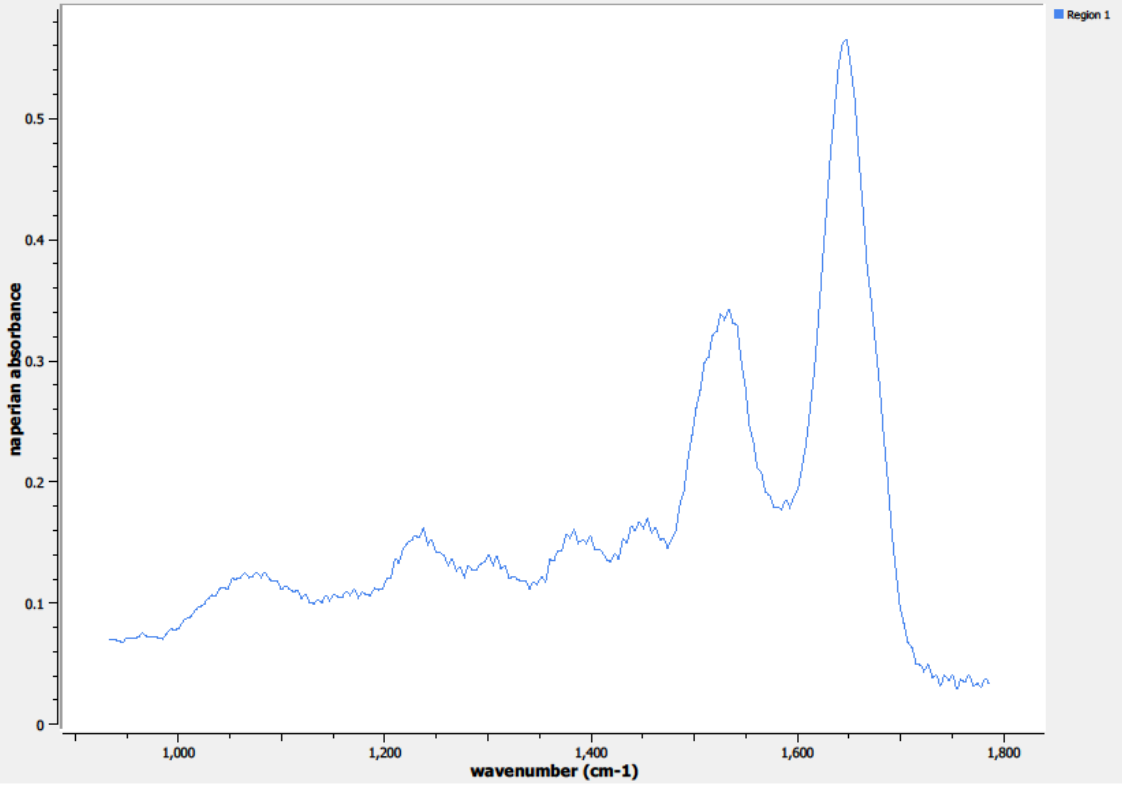



**Figure 0.90. Point P Spectra**

## A11. 91-10-A177

| | |
|---|---|
| Unique ID | 91-10-A177 |
| Age | 29 |
| Race | unknown |
| Location of Sample Collection | cervix |
| Notes of Interest | cervix tumor, papillary adenocarcinoma of cervix |

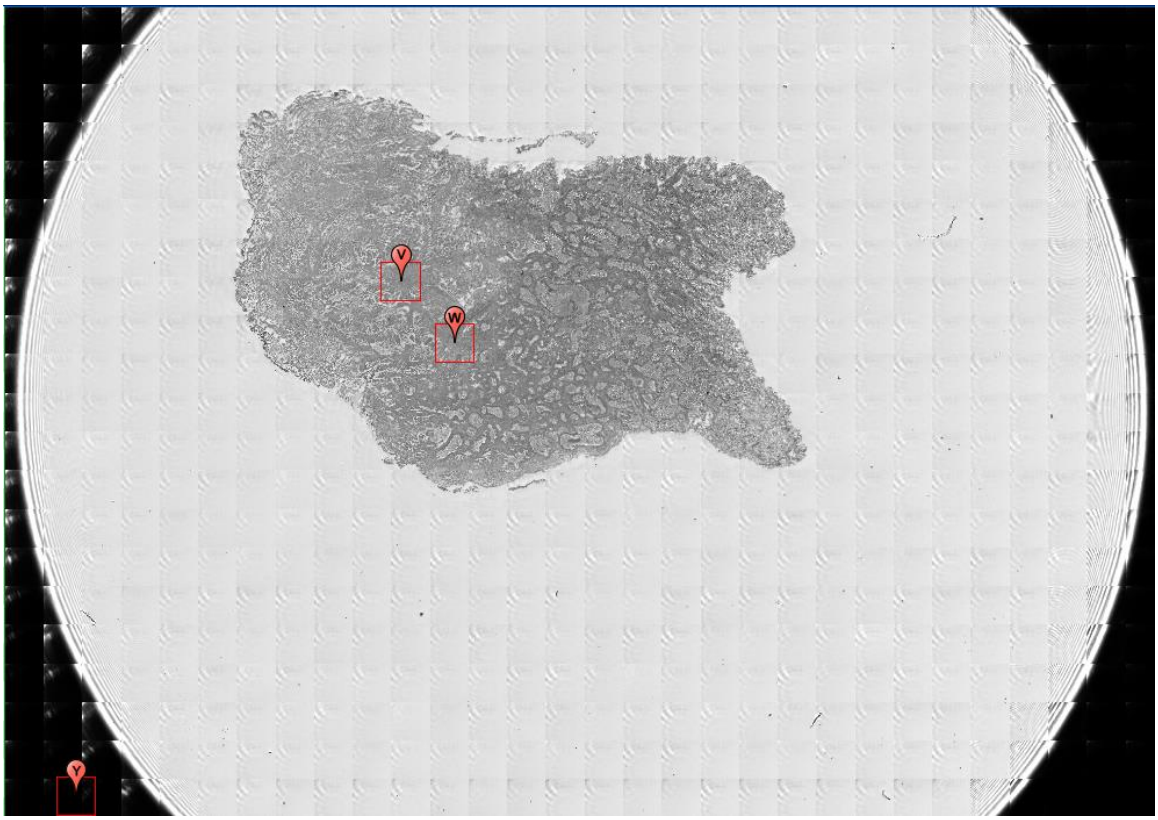


**Figure 0.91. Specimen 91-10-A177, cervical cancer tissue**

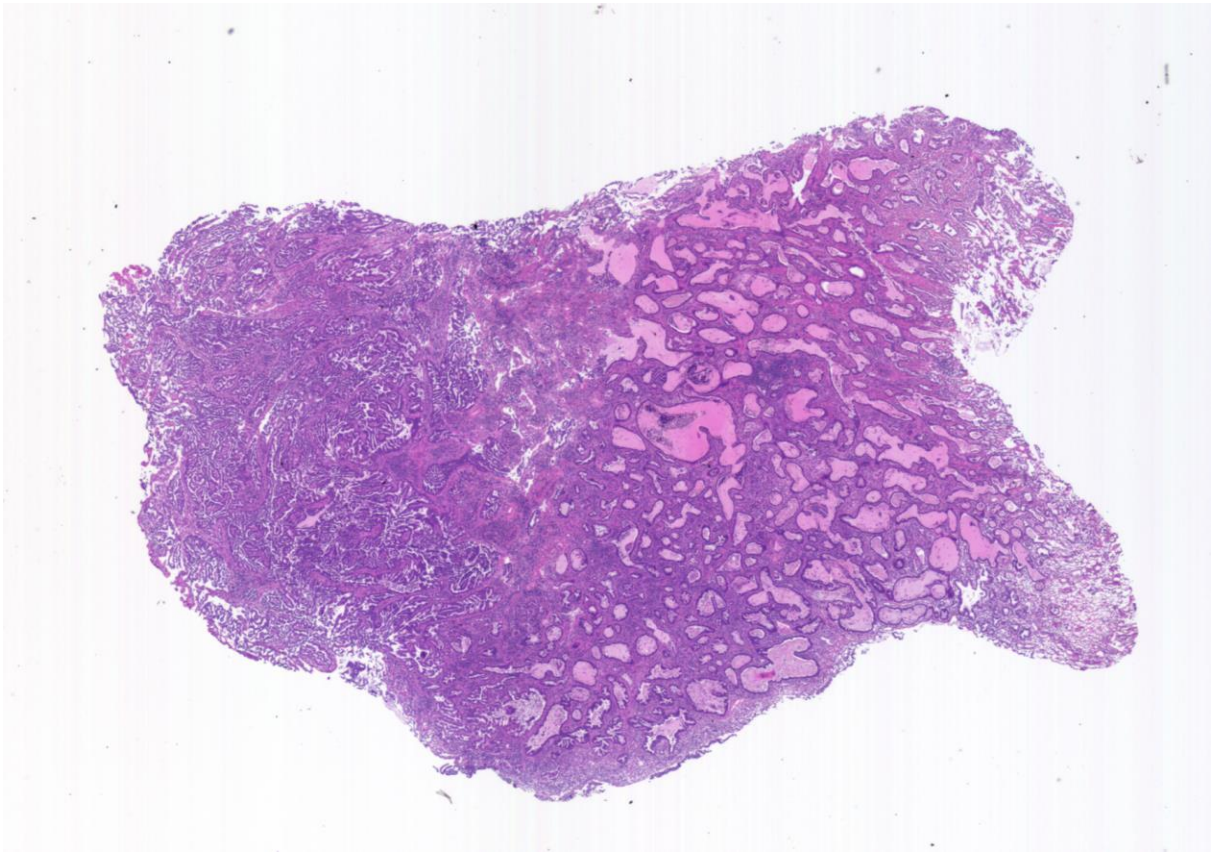


**Figure 0.92. Specimen 91-10-A177, H&E stain, papillary adenocarcinoma of the cervix**

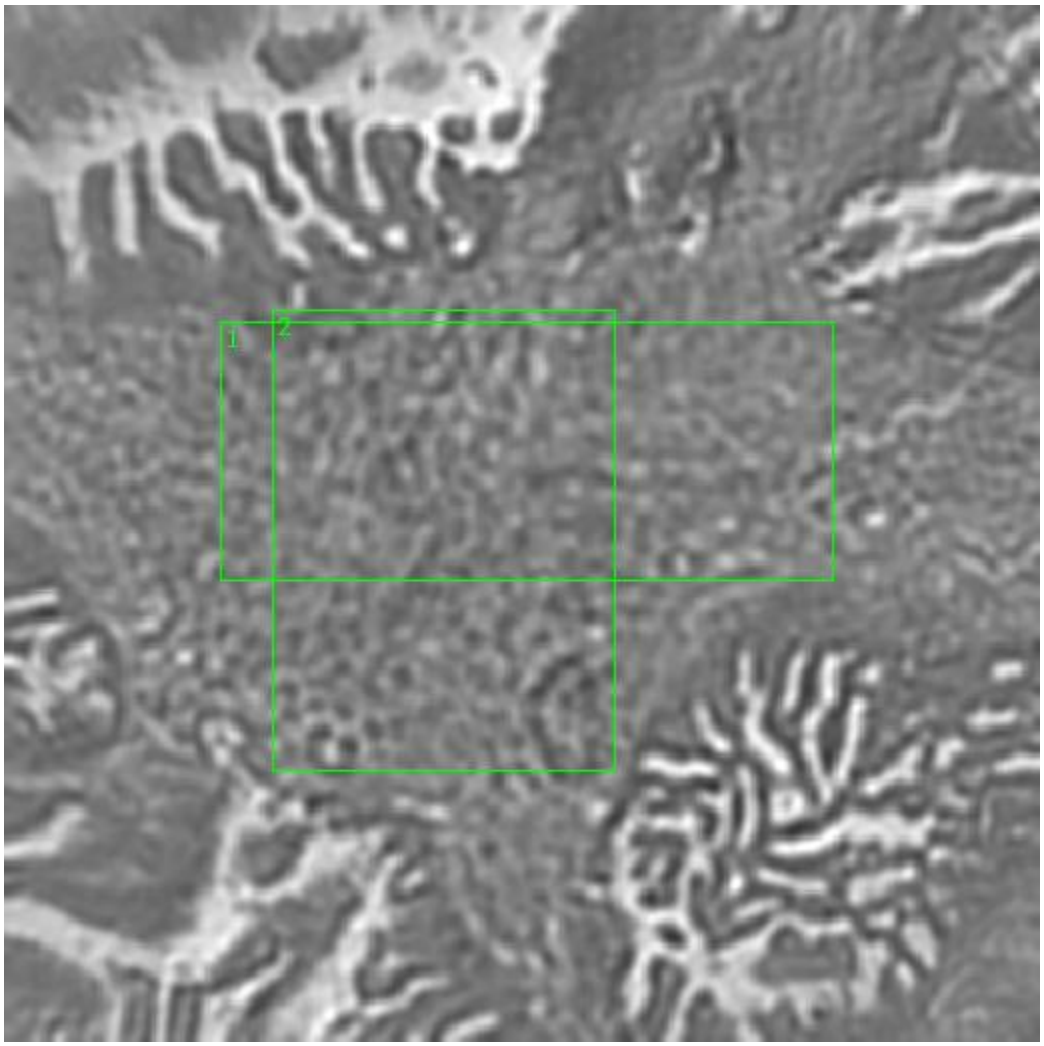



**Figure 0.93. Point V Regions of Interest**
**Regions 1, 2: mucous-secreting glands, papillary adenocarcinoma**

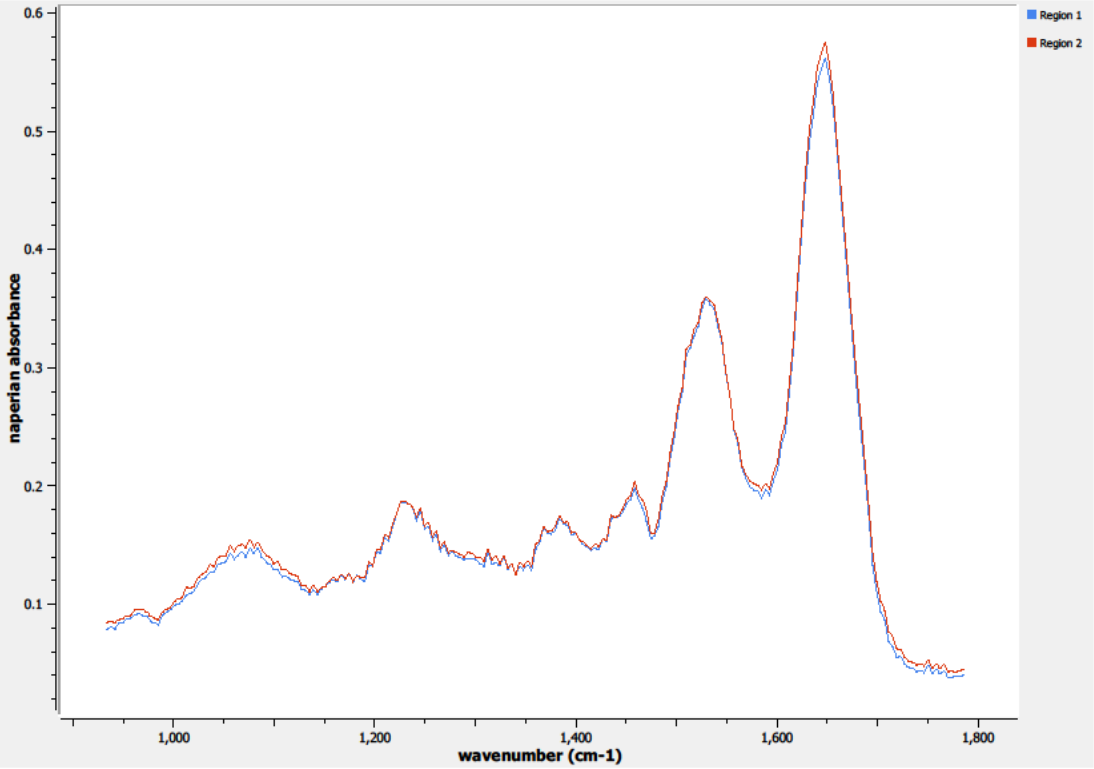



**Figure 0.94. Point V Spectra**

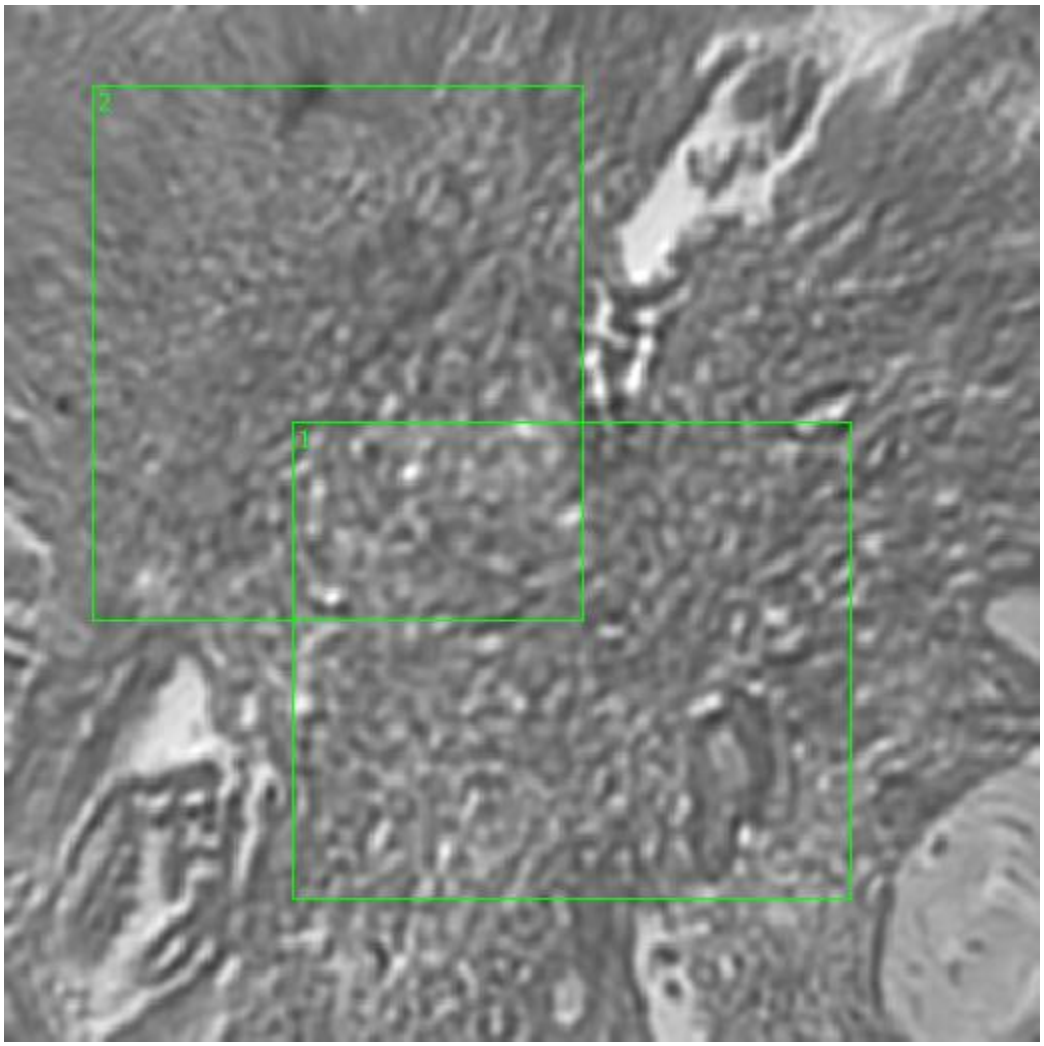



**Figure 0.95. Point W Regions of Interest**
**Regions 1, 2: mucous-secreting glands, papillary adenocarcinoma**

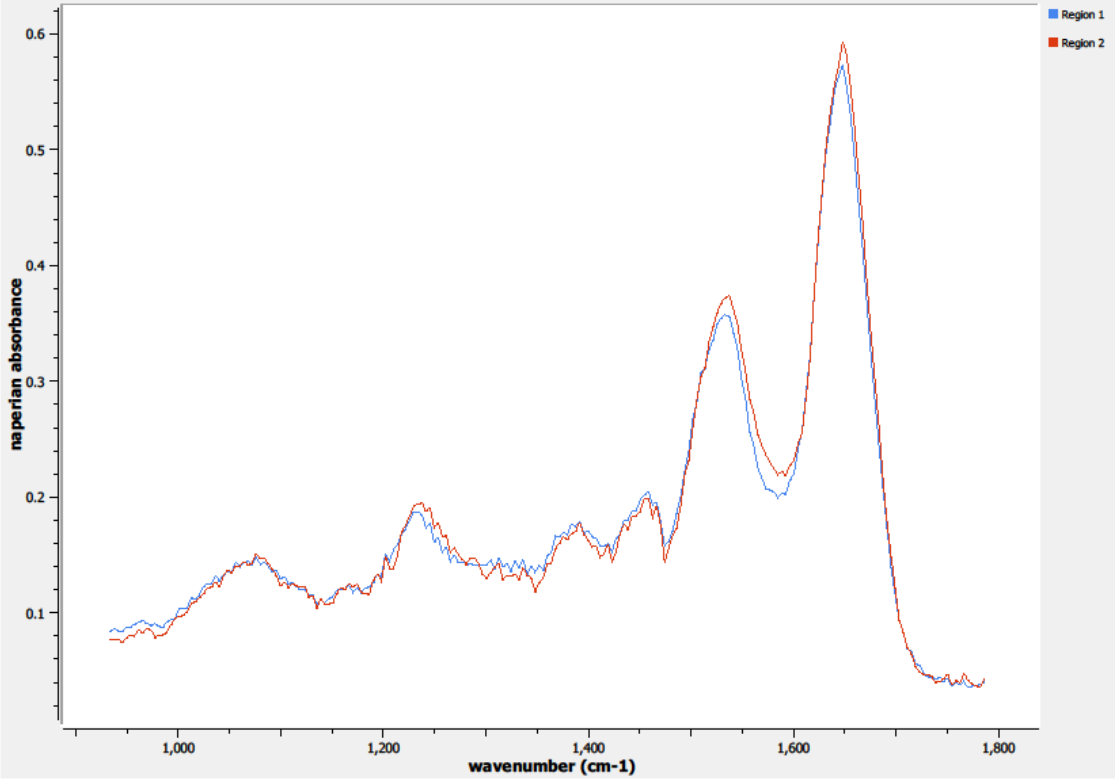



**Figure 0.96. Point W Spectra**

## A12. 92-12-A198

| | |
|---|---|
| Unique ID | 92-12-A198 |
| Age | 61 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | squamous cell carcinoma, poorly differentiated, fixed in OCT, unused in data results |

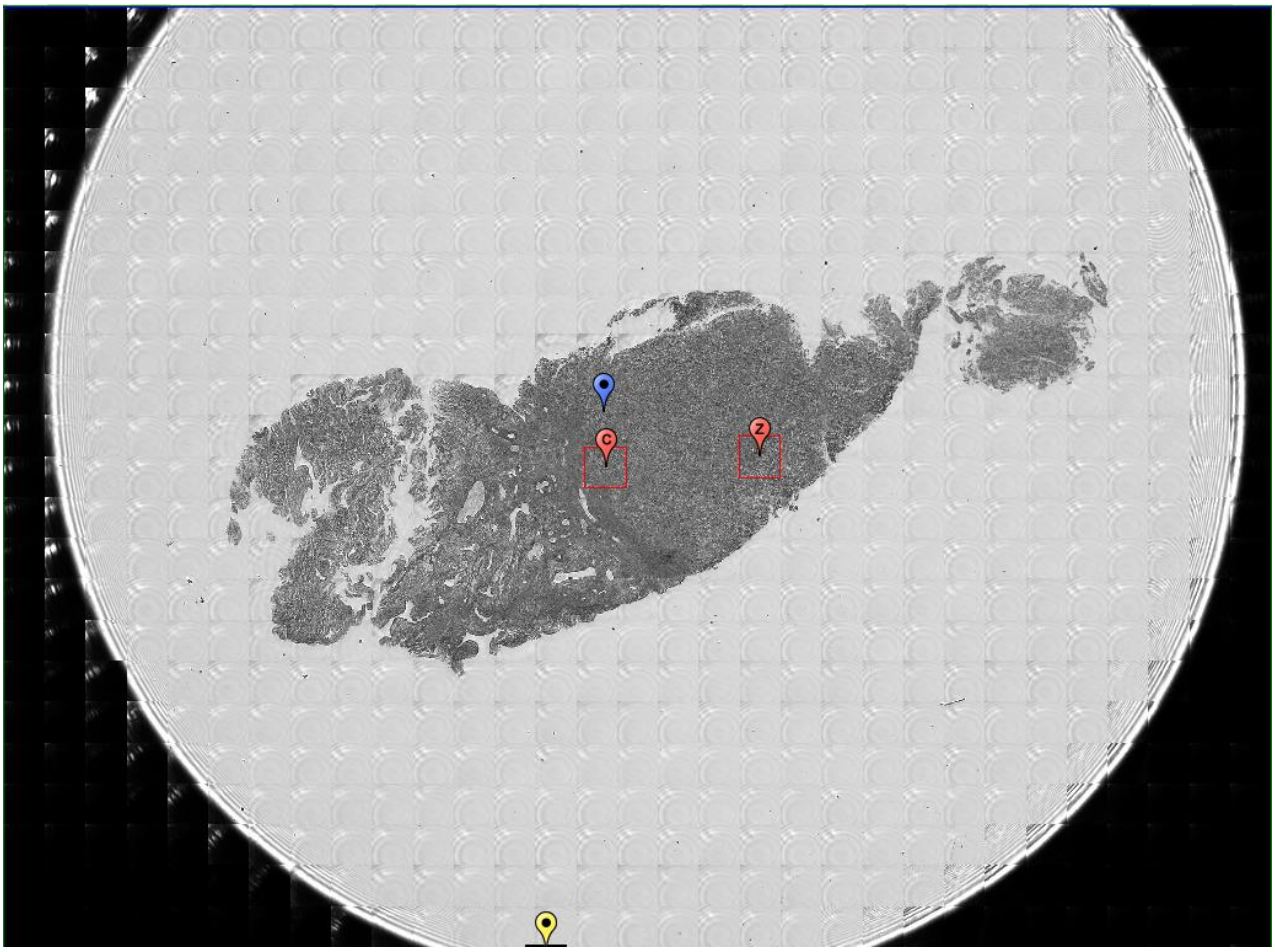



**Figure 0.97. Specimen 92-12-A198, cervical cancer tissue**

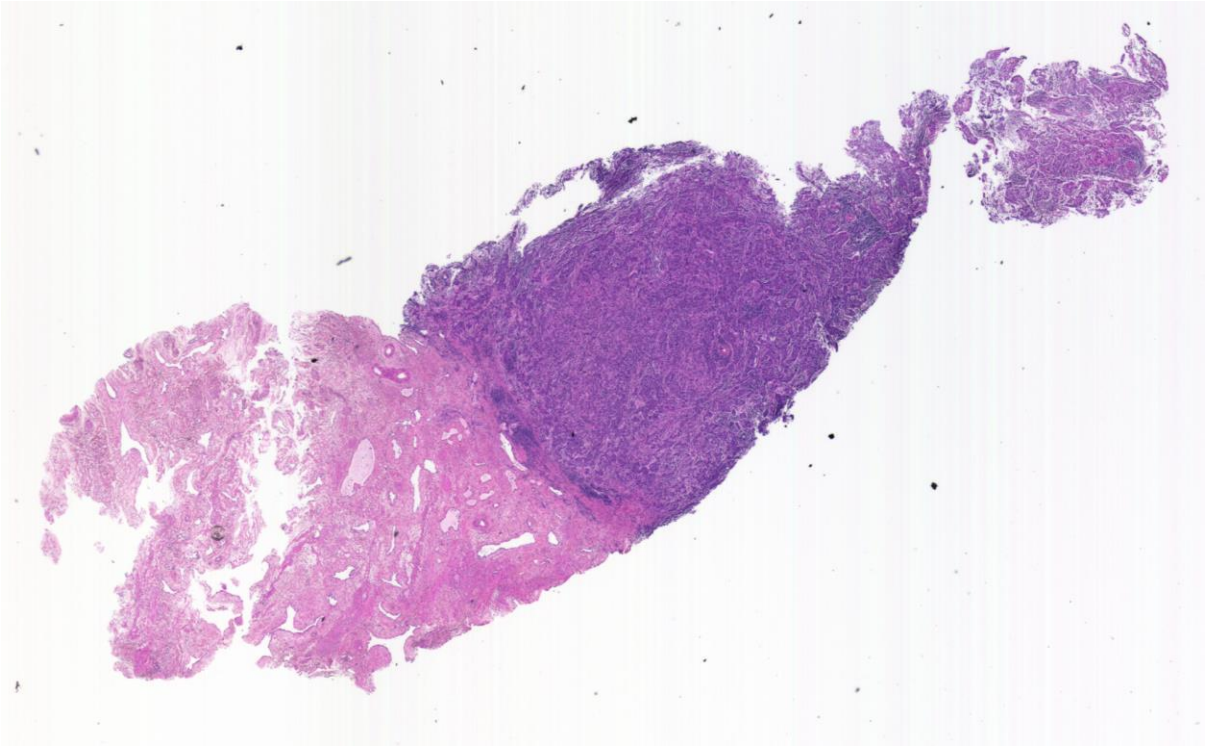


**Figure 0.98. Specimen 92-12-A198, H&E stain, squamous cell carcinoma of the cervix**

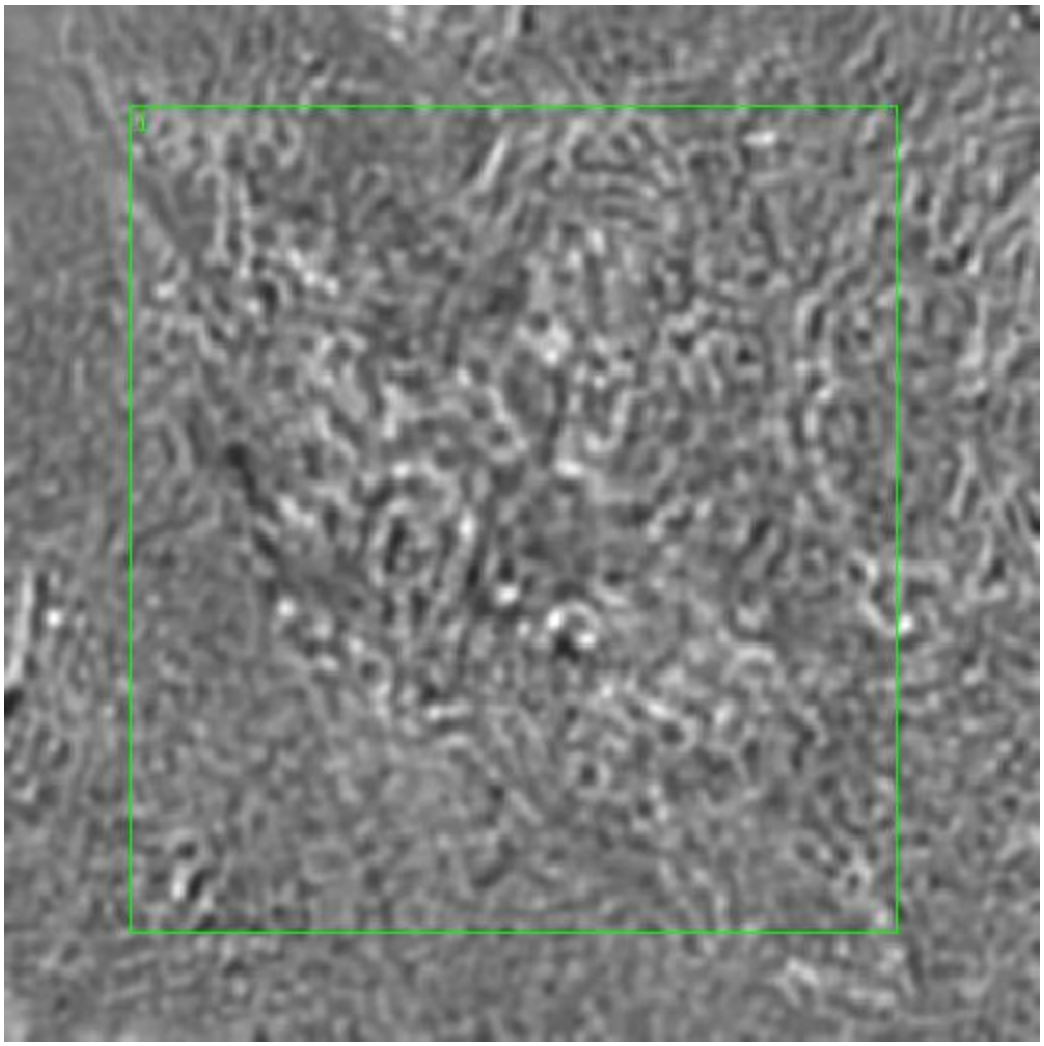


**Figure 0.99. Point C Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

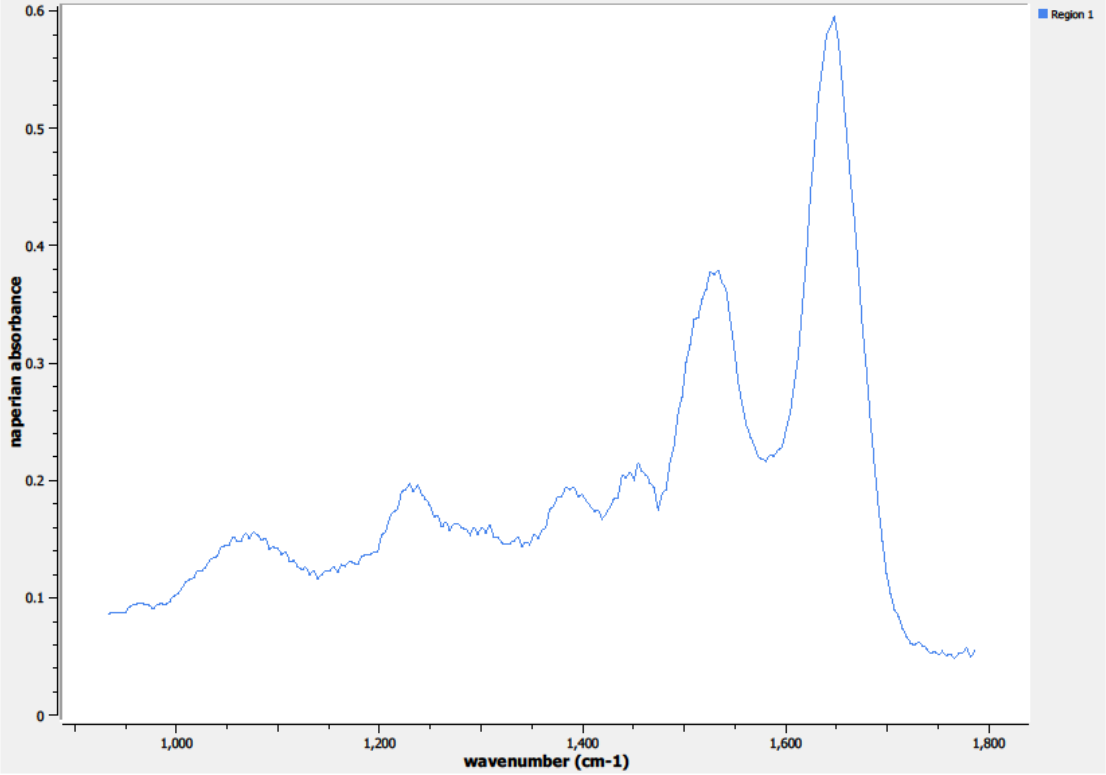



**Figure 0.100. Point C Spectra**

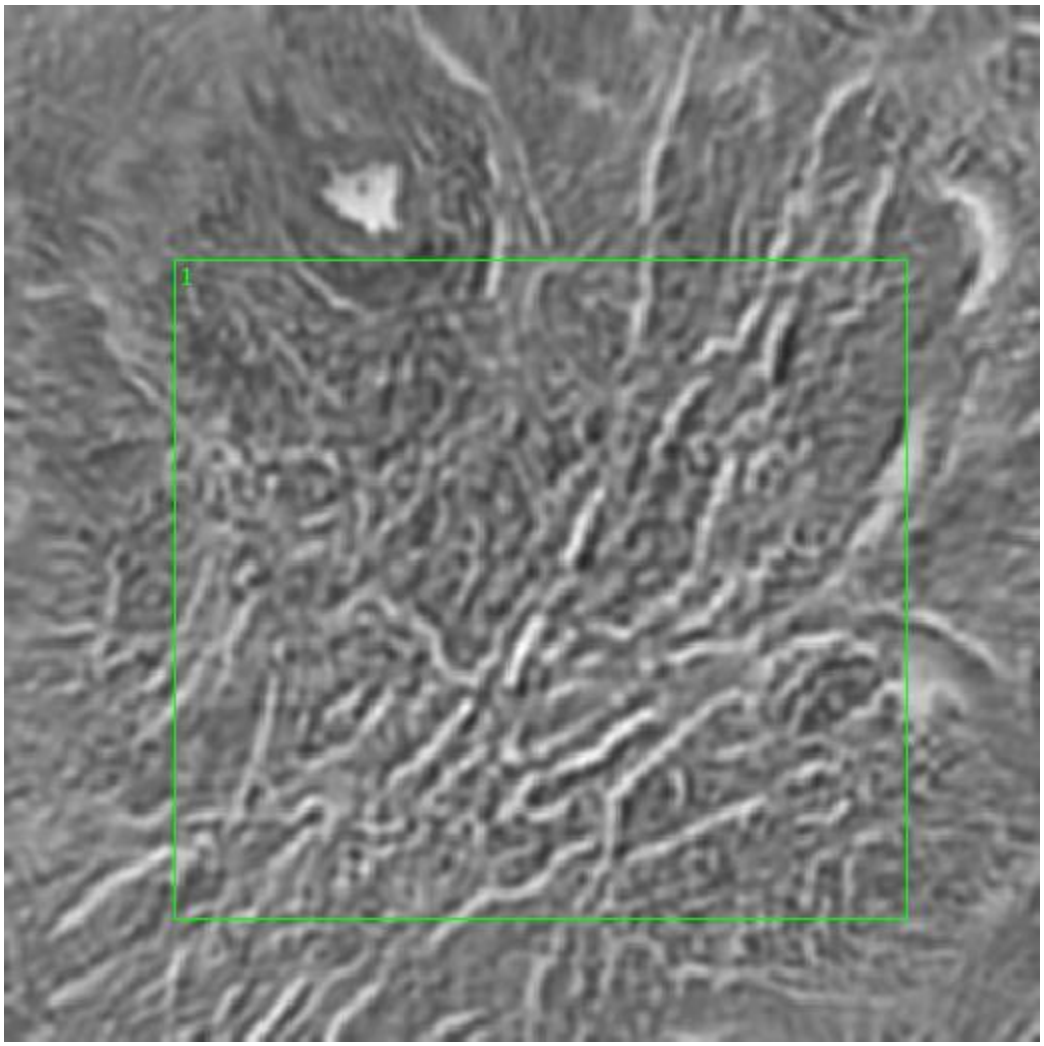



**Figure 0.101. Point Z Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

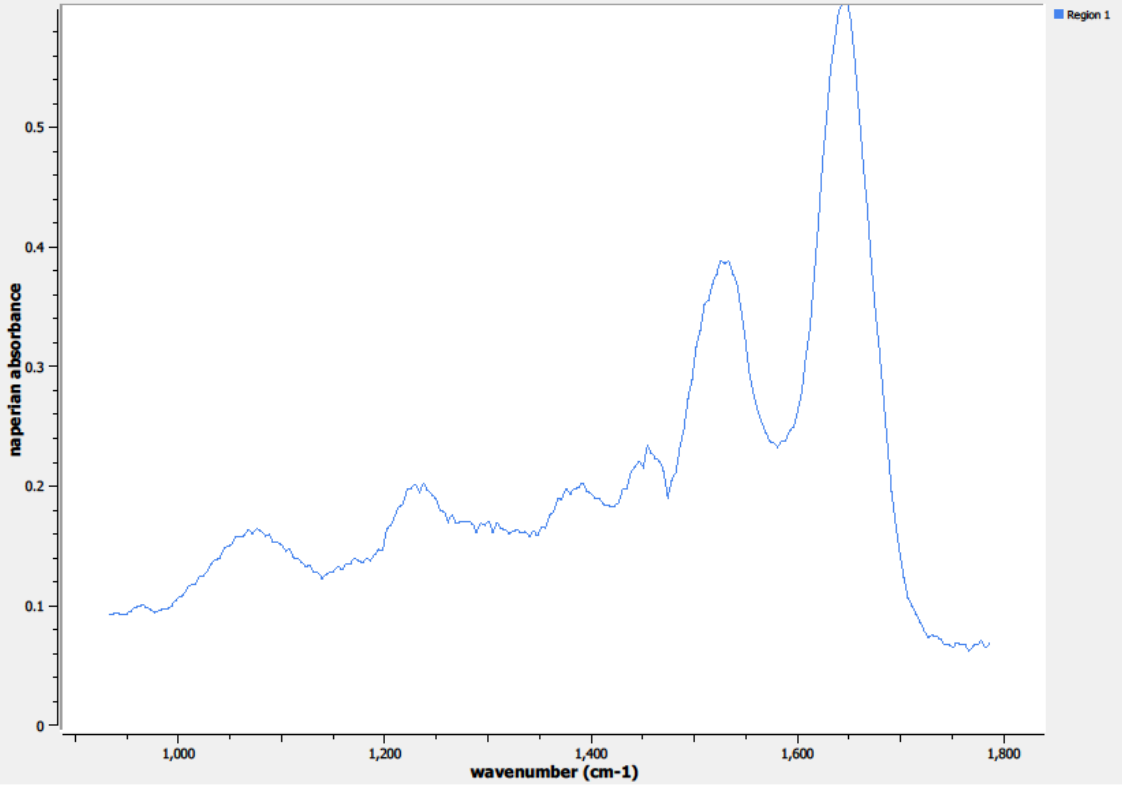



**Figure 0.102. Point Z Spectra**

## A13. 93-02-A251

| | |
|---|---|
| Unique ID | 93-02-A251 |
| Age | 22 |
| Race | white |
| Location of Sample Collection | cervix |
| Notes of Interest | squamous cell carcinoma |

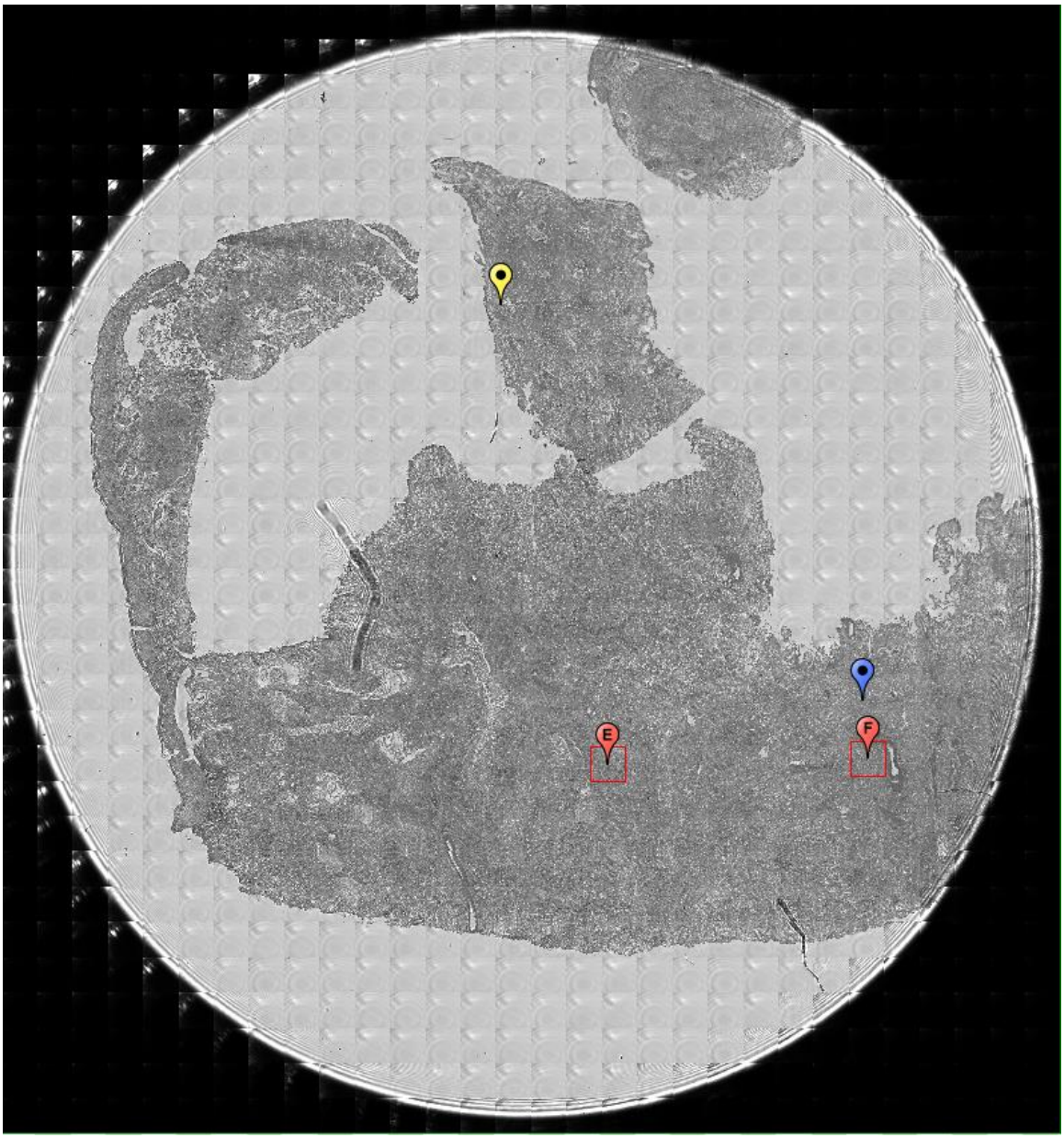



**Figure 0.103. Specimen 93-02-A251, cervical cancer tissue**

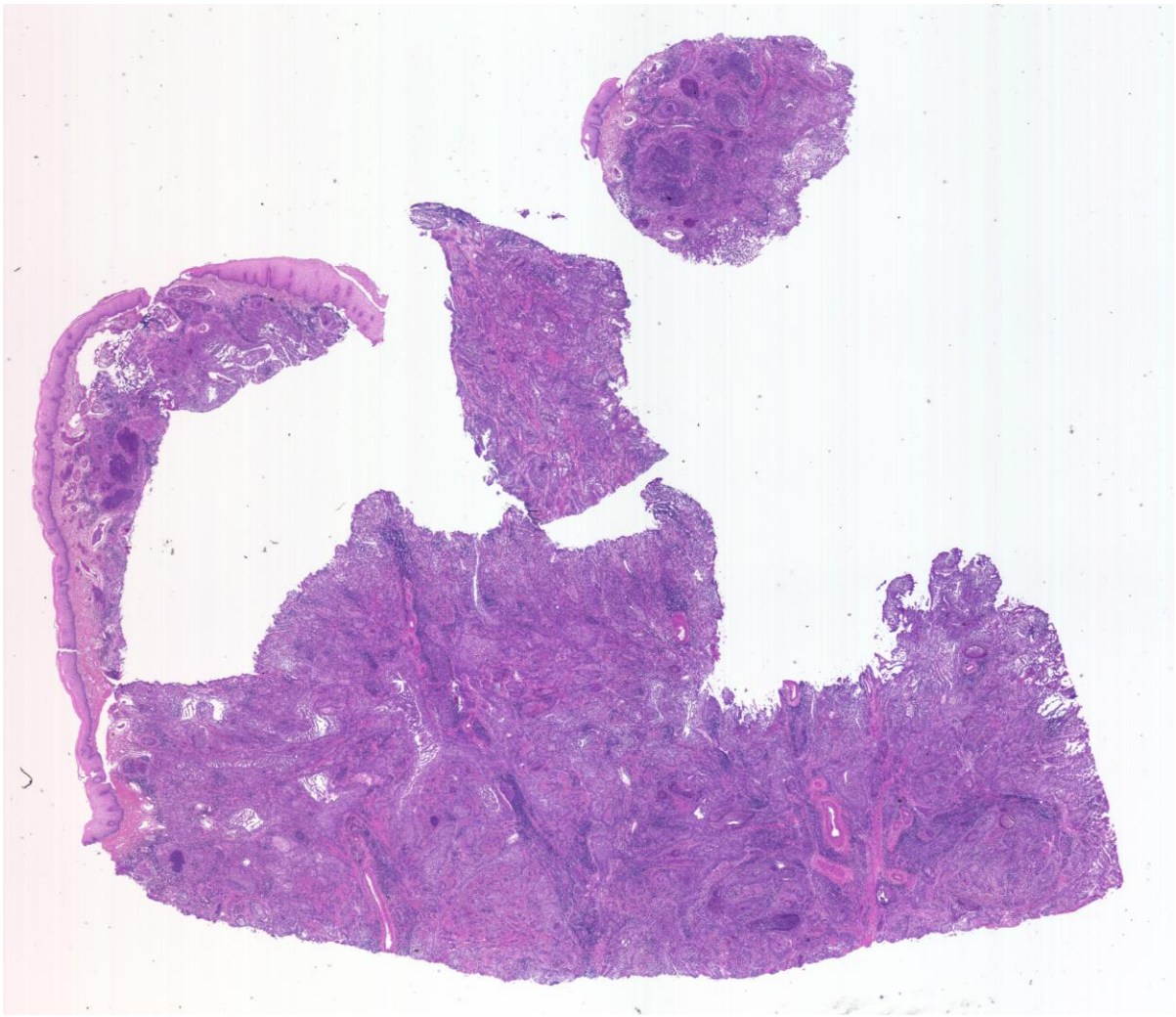


**Figure 0.104. Specimen 93-02-A251, H&E stain, squamous cell carcinoma of the cervix**

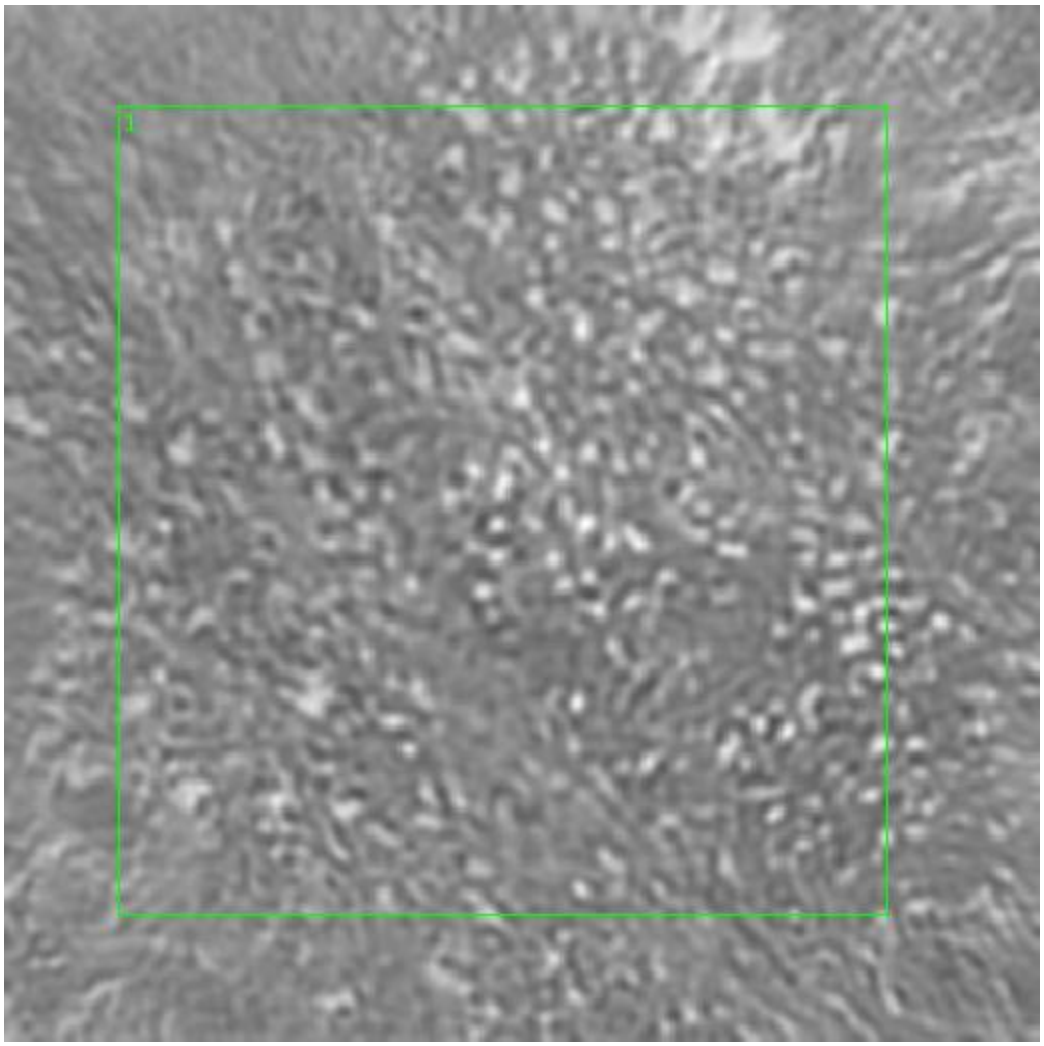


**Figure 0.105. Point E Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

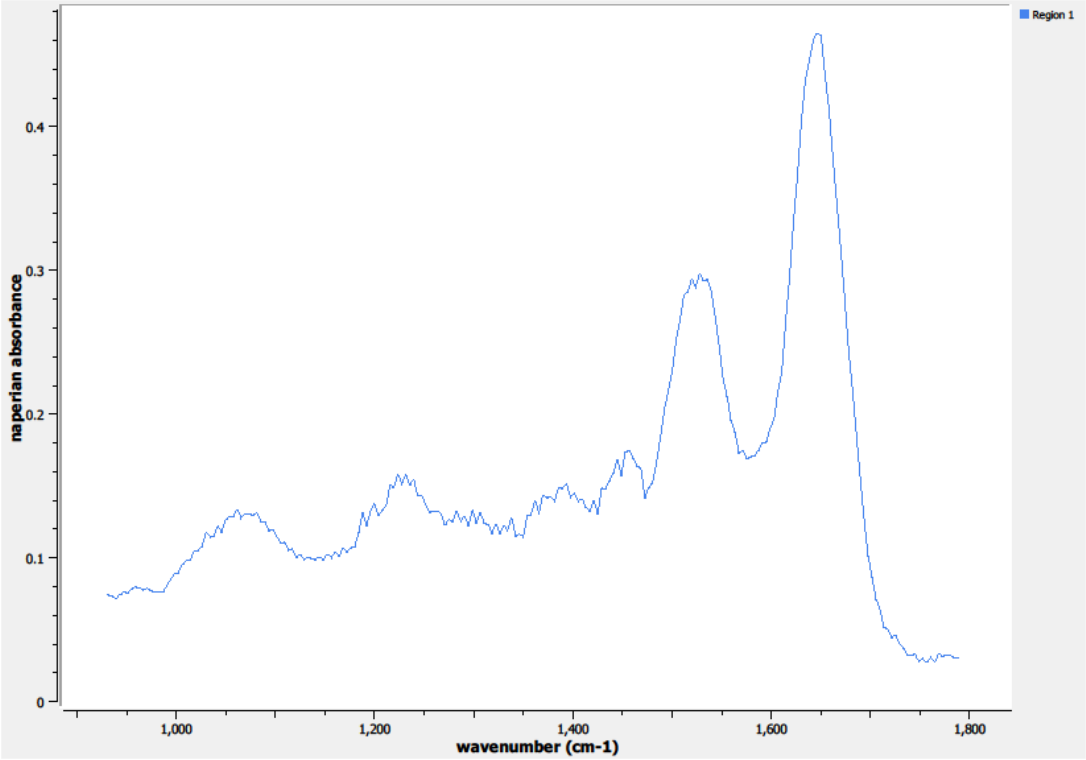



**Figure 0.106. Point E Spectra**

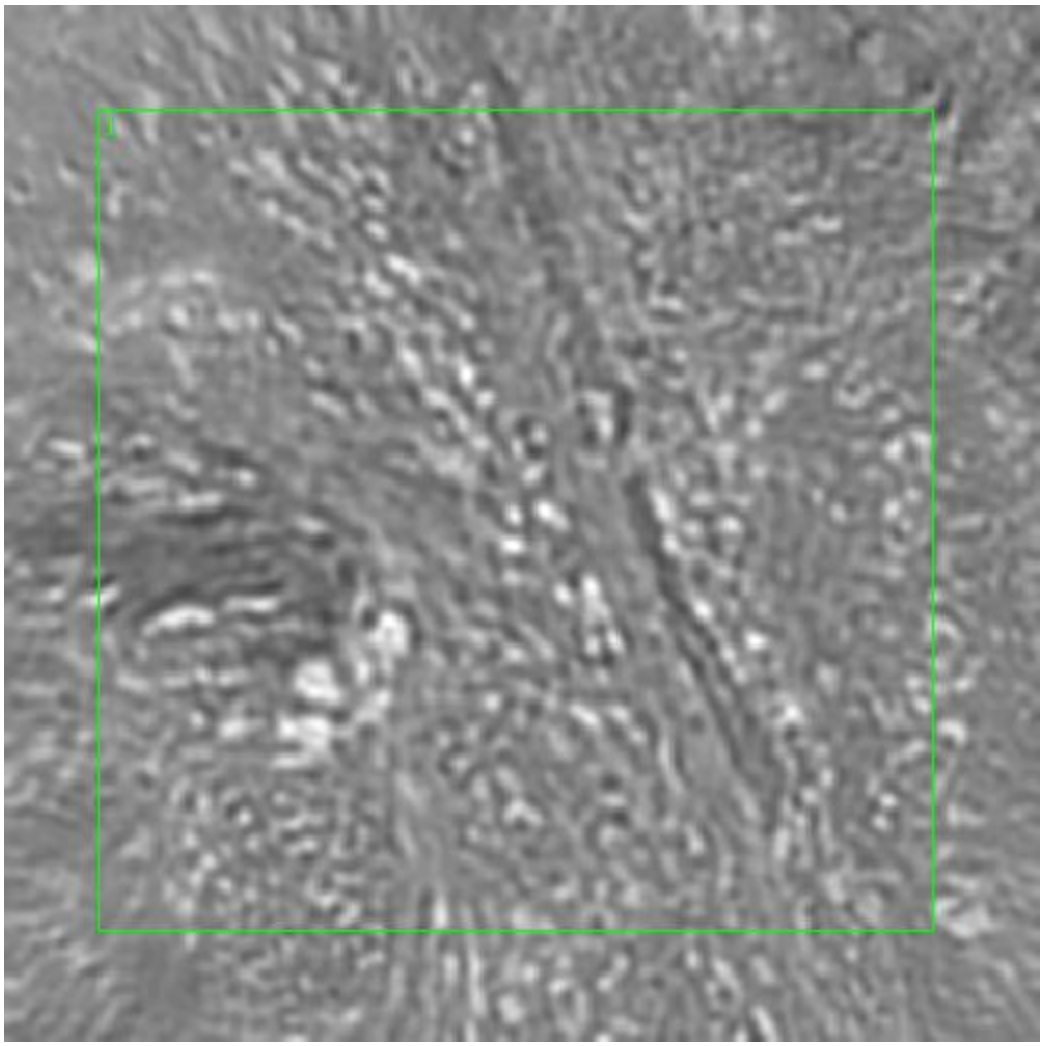


**Figure 0.107. Point F Regions of Interest**
**Region 1: lamina propia, squamous cell carcinoma**

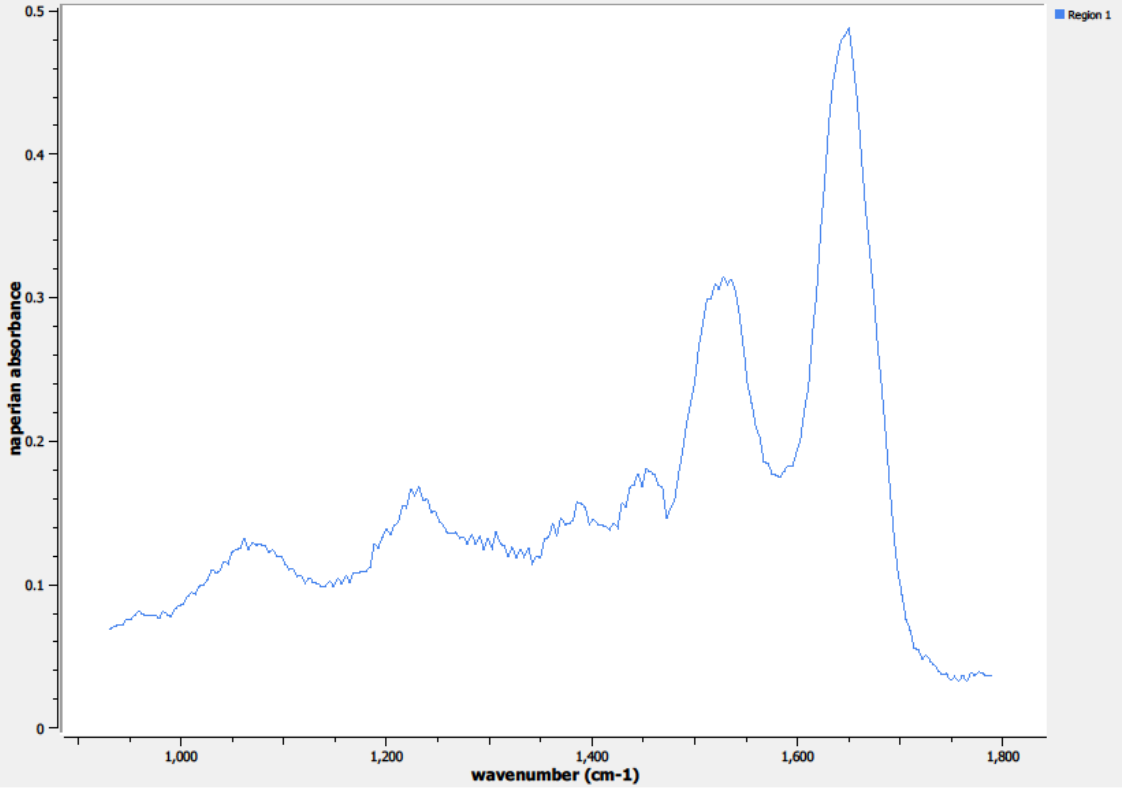



**Figure 0.108. Point F Spectra**

# Appendix C

# Breast Normal

**A1.10-09-A133d**

| | |
|---|---|
| Unique ID | 10-09-A133d |
| Age | 57 |
| Race | white |
| Location of Sample Collection | left breast |
| Notes of Interest | breast reduction, damaged slide, unused in data results |

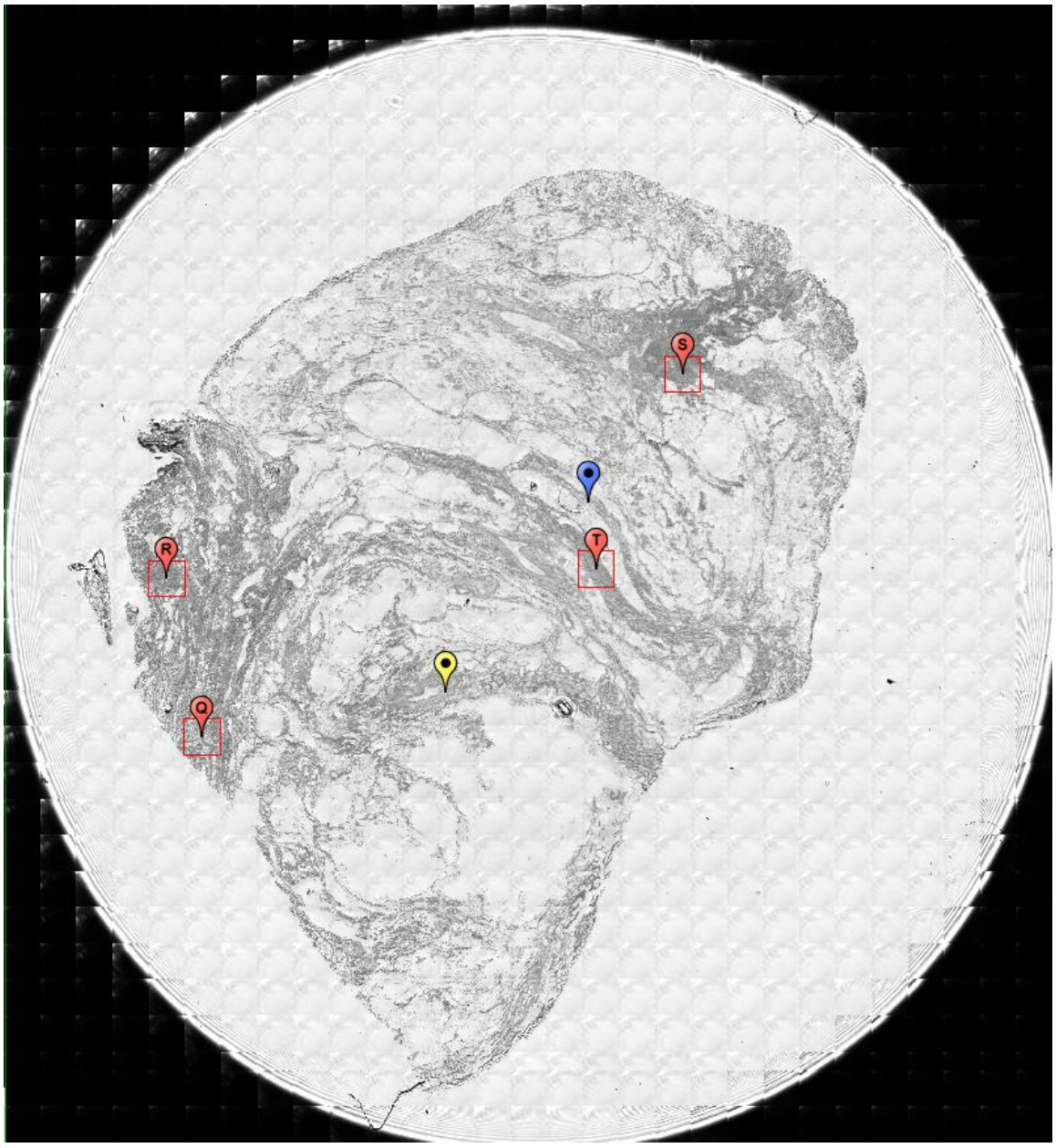



**Figure 0.1. Specimen 10-09-A133d, normal left breast tissue.**

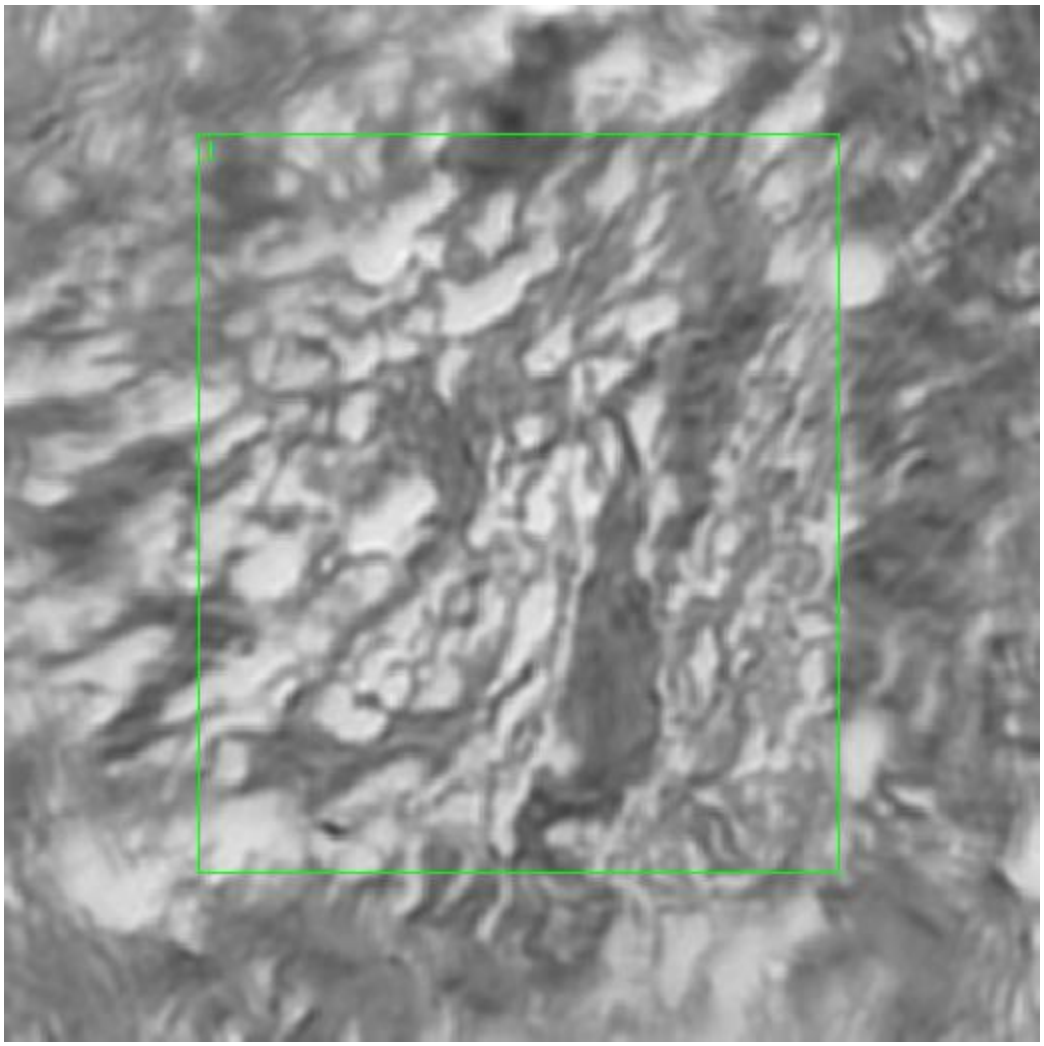


**Figure 0.2. Point Q Regions of Interest**
**Region 1: left breast**

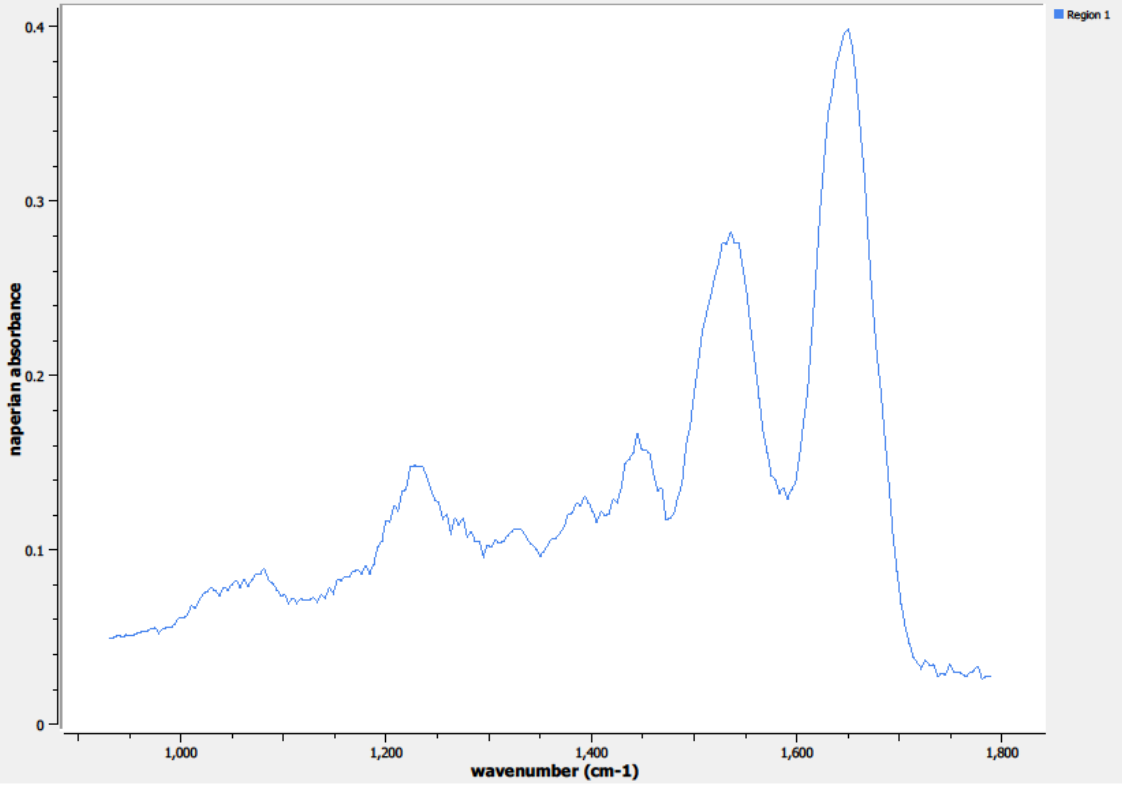



**Figure 0.3. Point Q Spectra**

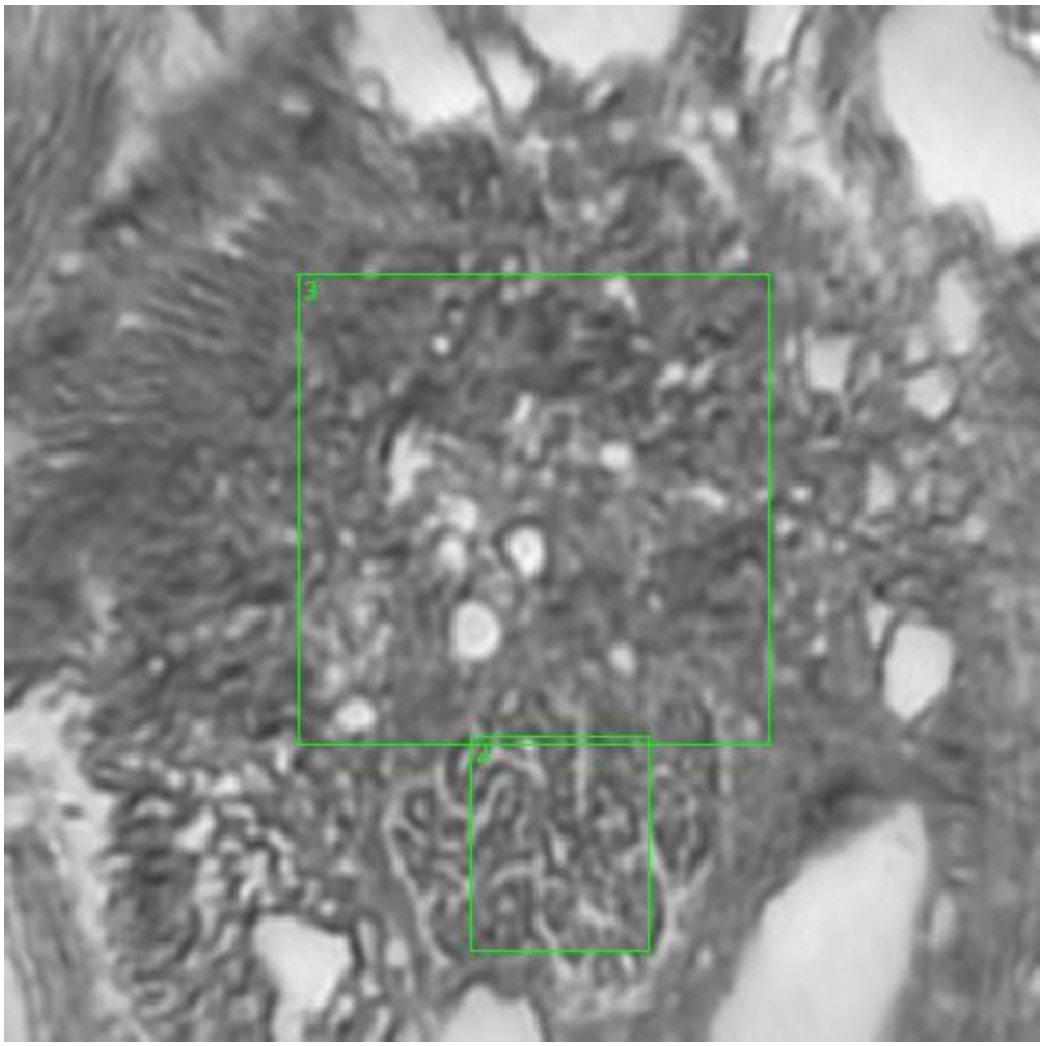



**Figure 0.4. Point R Regions of Interest**
**Regions 2, 3: left breast**

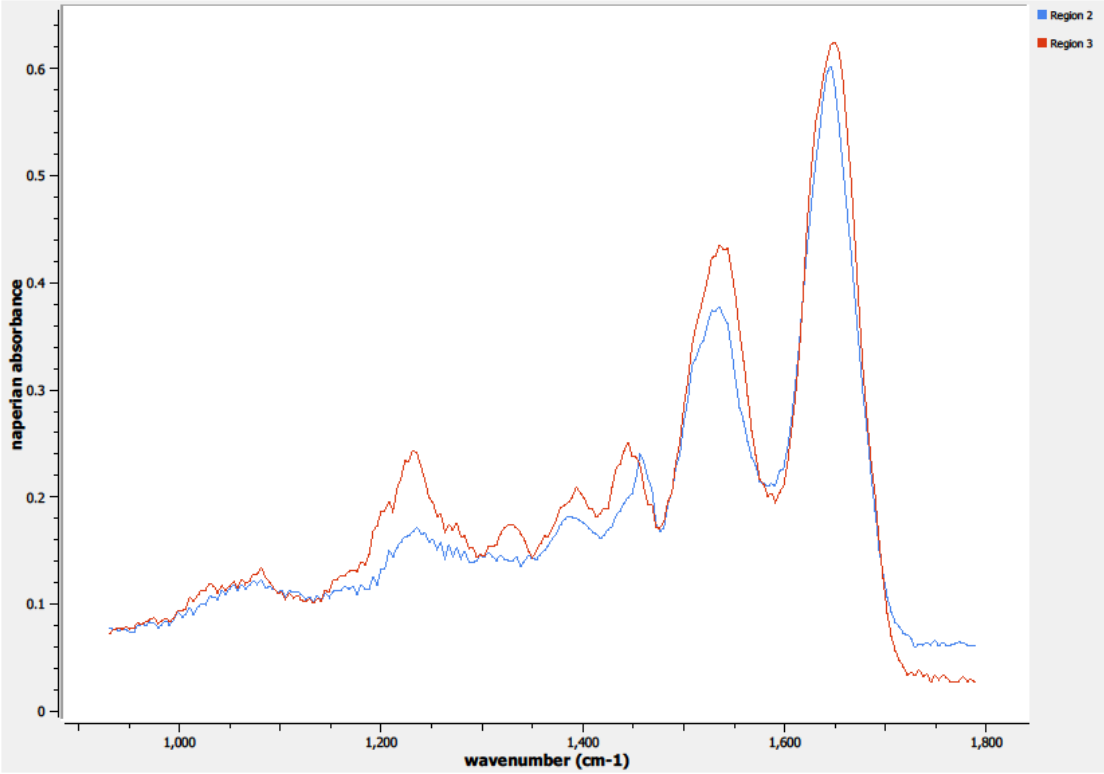



**Figure 0.5. Point R Spectra**

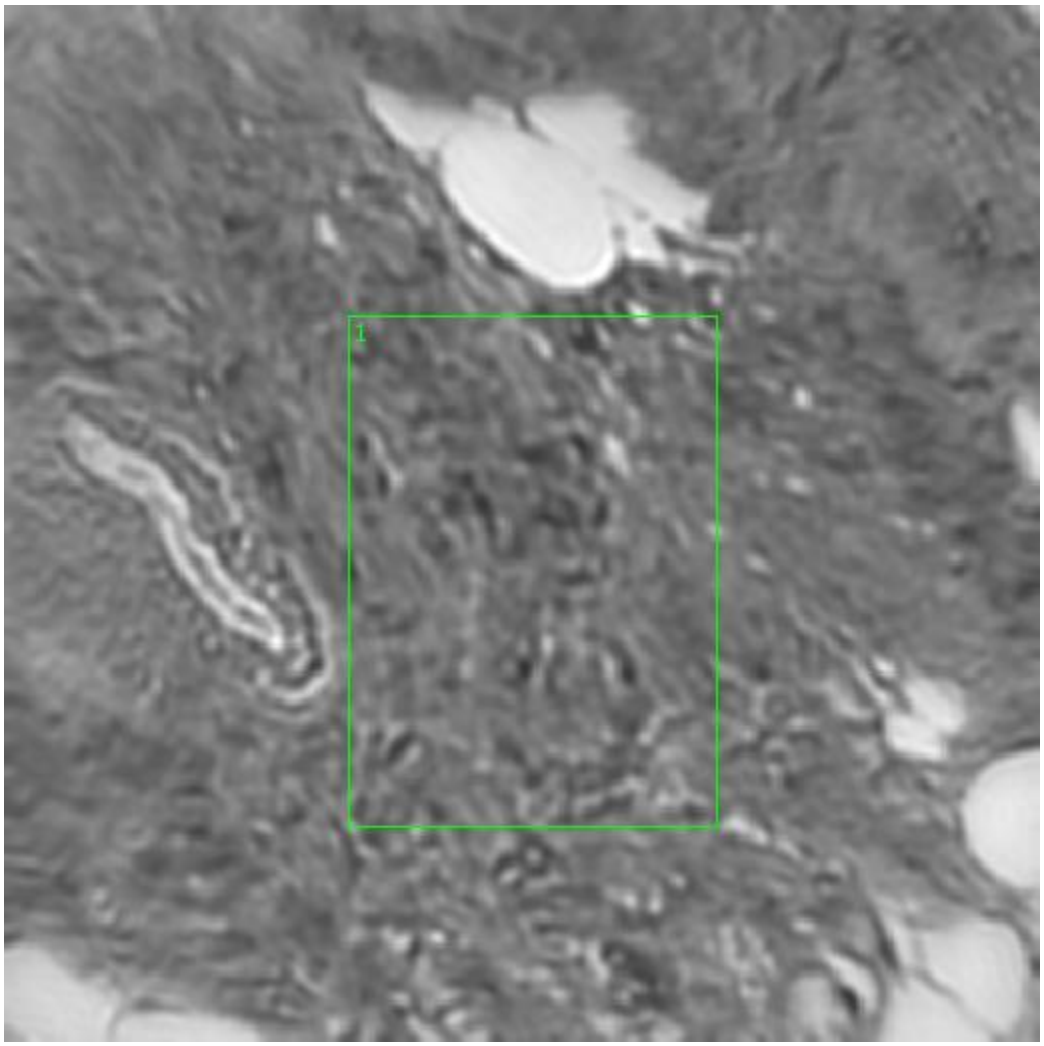



**Figure 0.6. Point S Regions of Interest**
**Region 1: left breast**

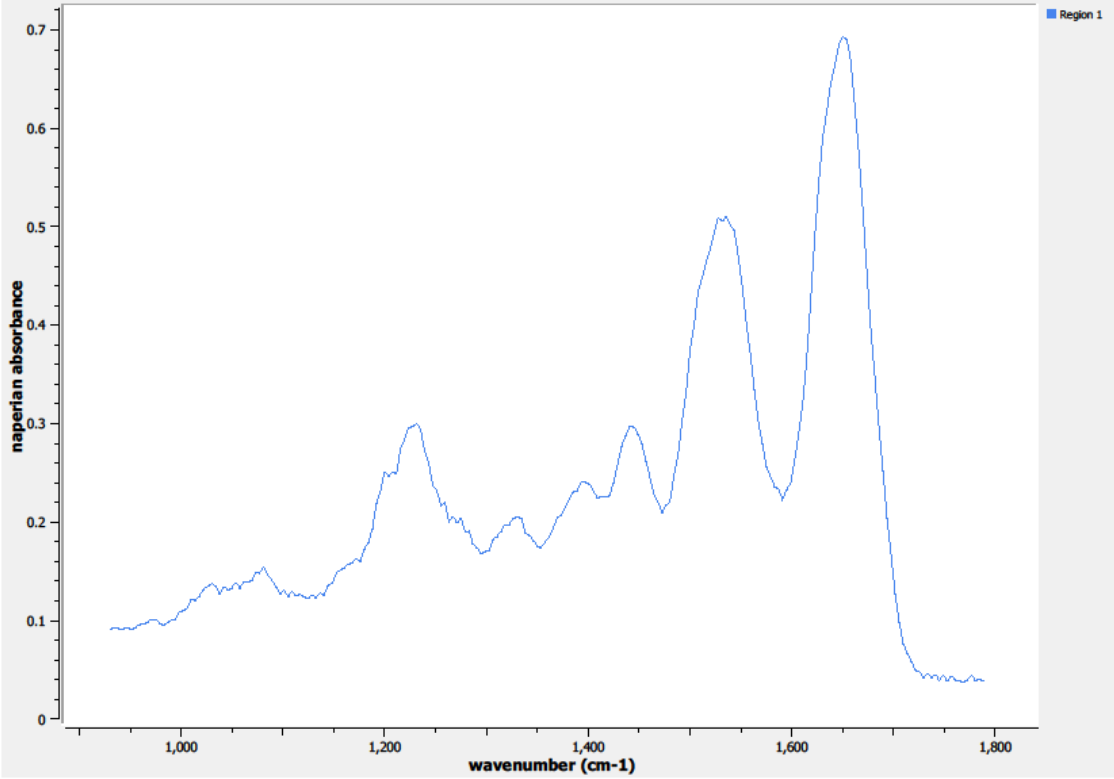



**Figure 0.7. Point S Spectra**

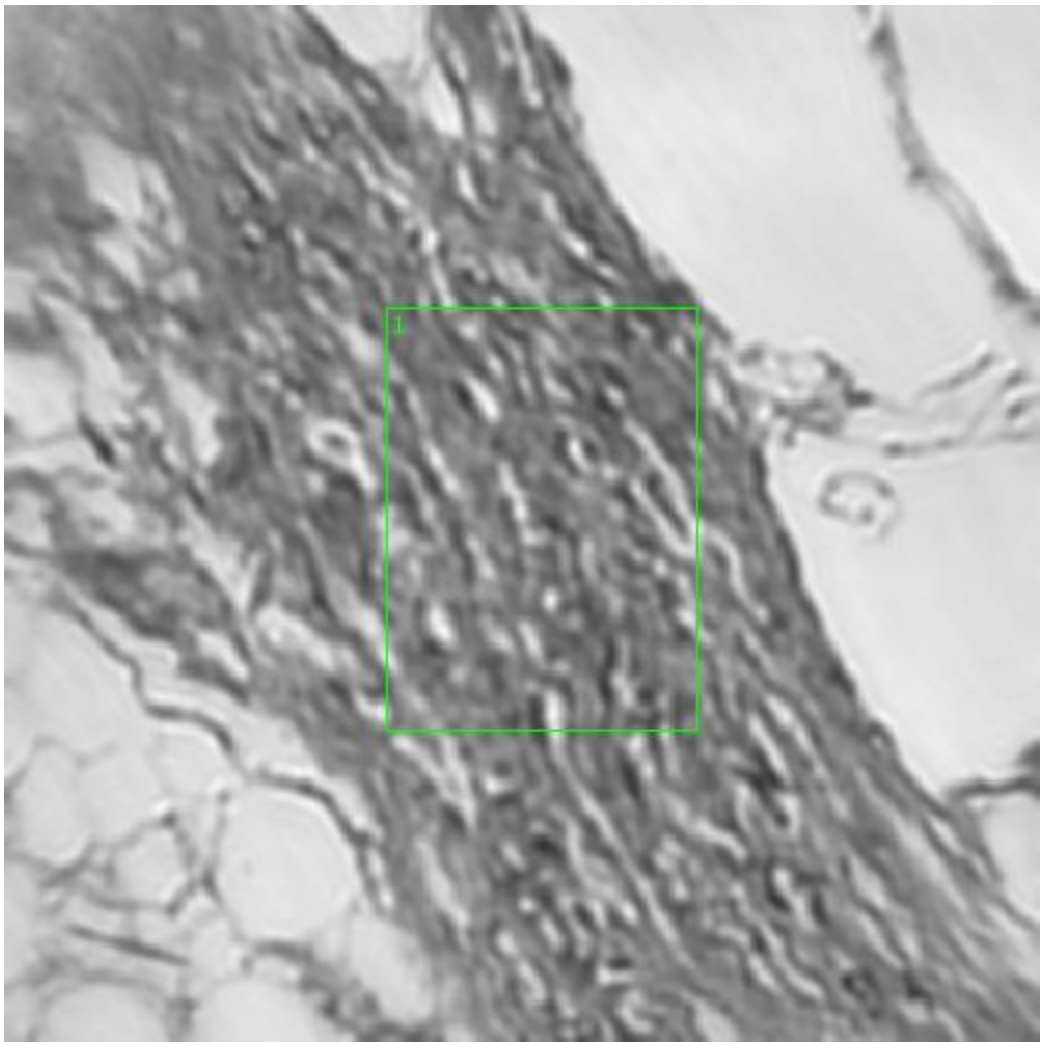


**Figure 0.8. Point T Regions of Interest**
**Region 1: left breast**

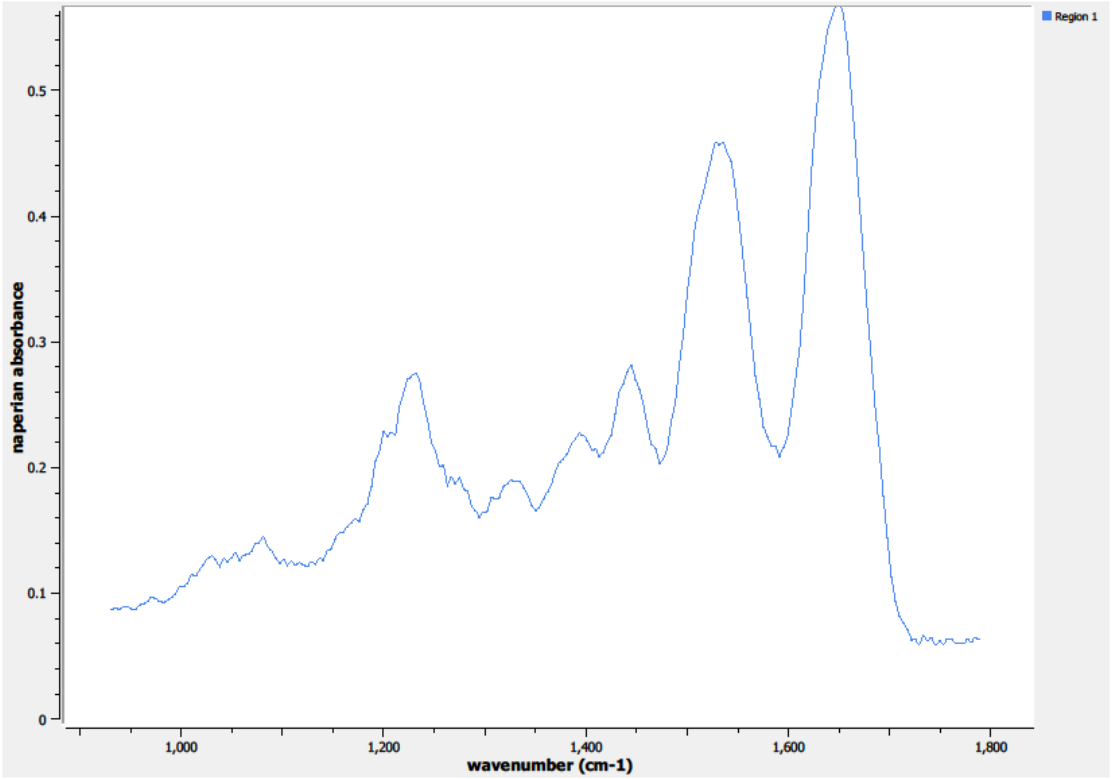



**Figure 0.9. Point T Spectra**

## A2.11-08-A085c

| | |
|---|---|
| Unique ID | 11-08-A085c |
| Age | 31 |
| Race | white |
| Location of Sample Collection | right breast |
| Notes of Interest | none |

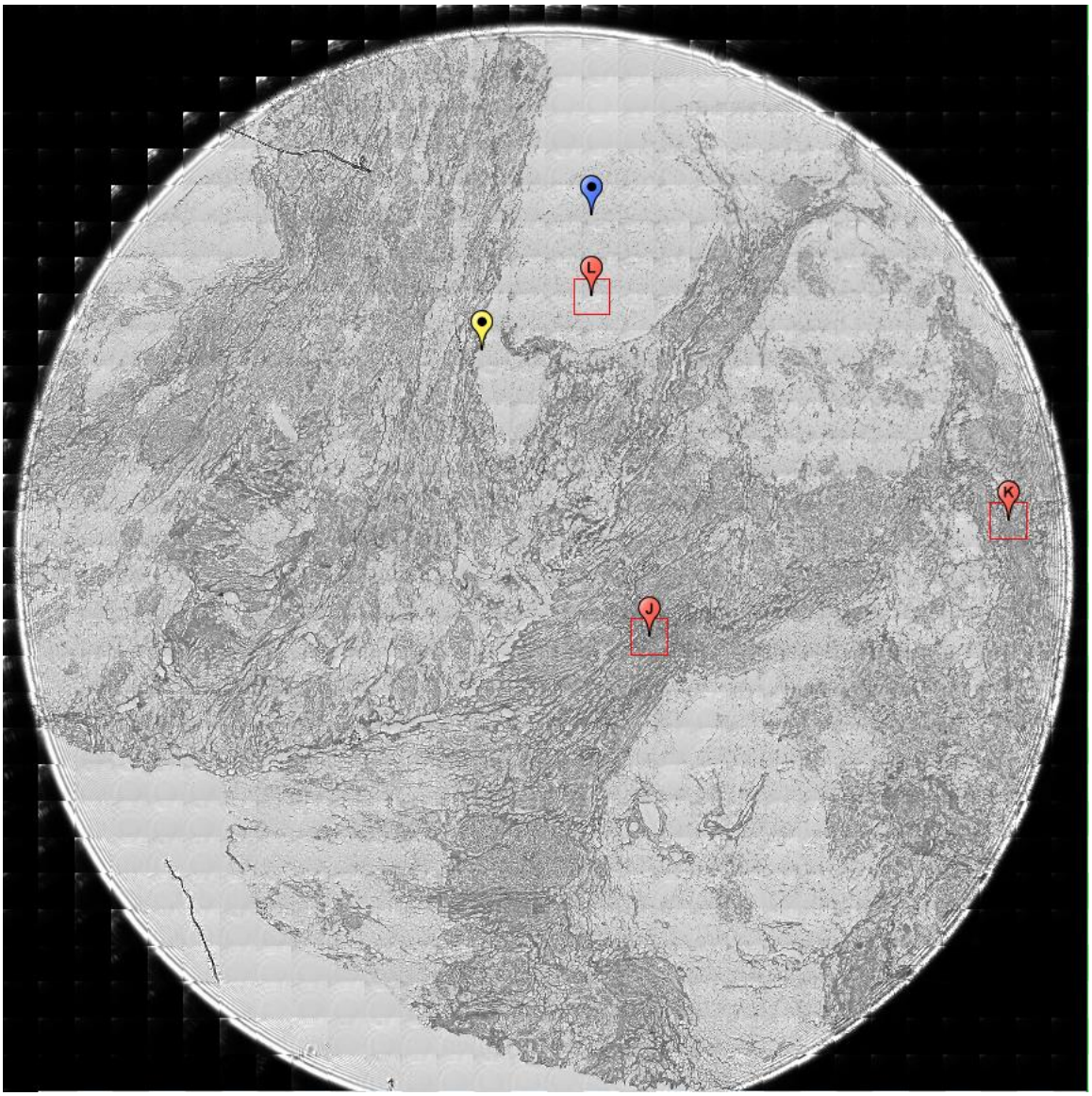



**Figure 0.10. Specimen 11-08-A085c, normal right breast tissue.**

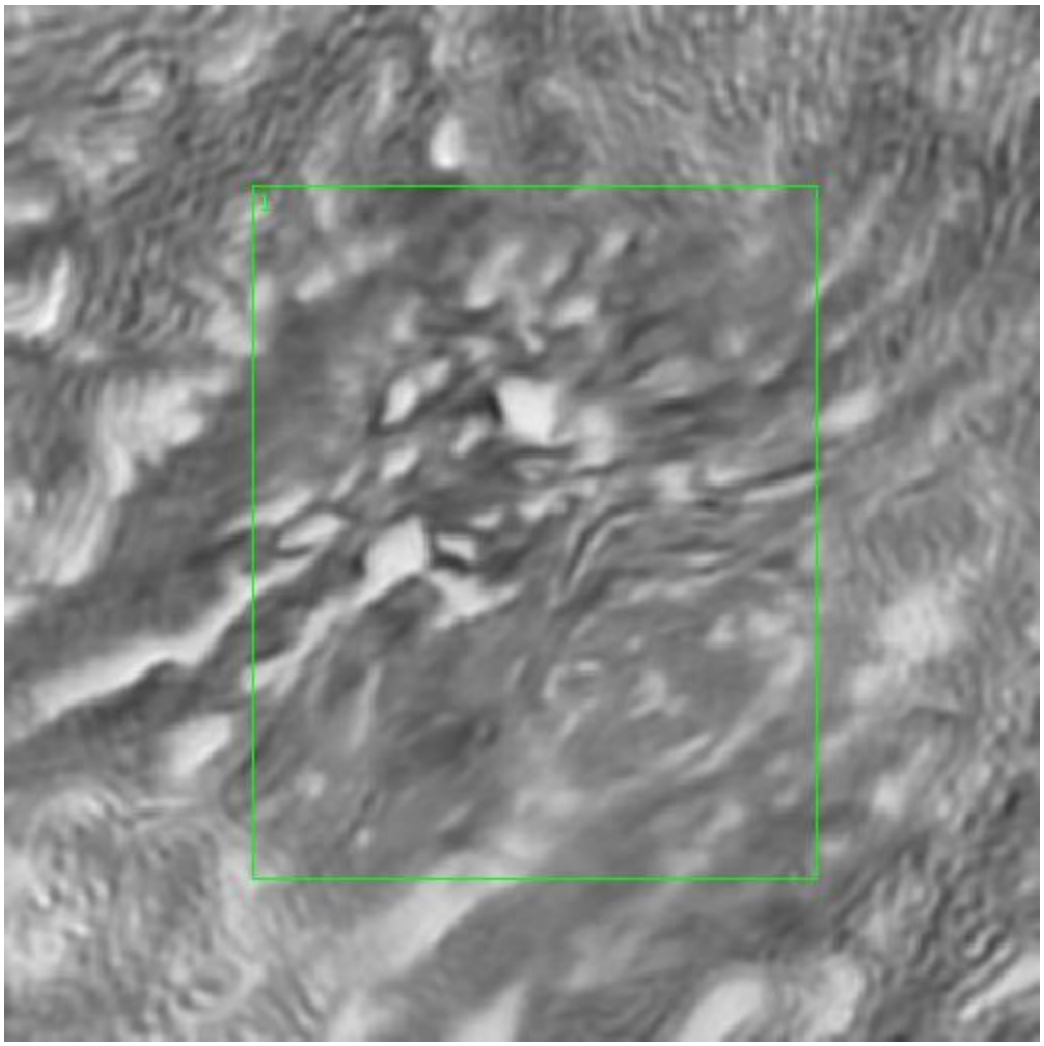


**Figure 0.11. Point J Regions of Interest**
**Region 1: normal right breast tissue**

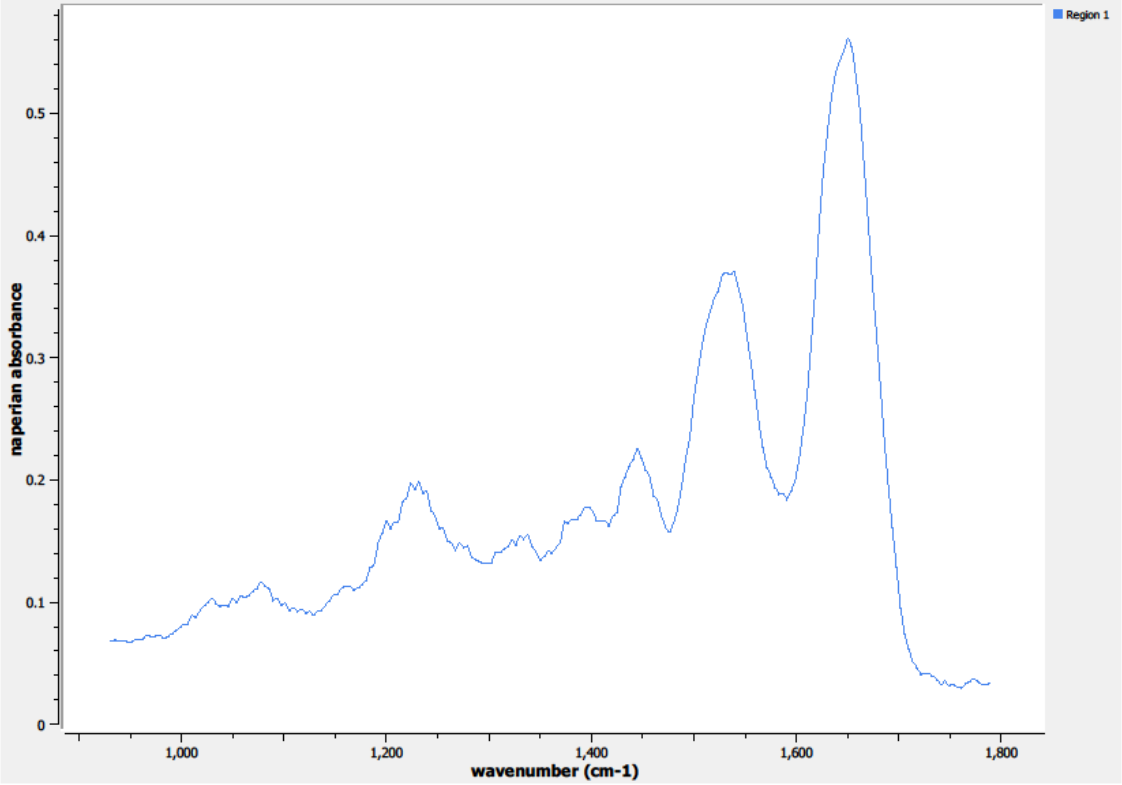



**Figure 0.12. Point J Spectra**

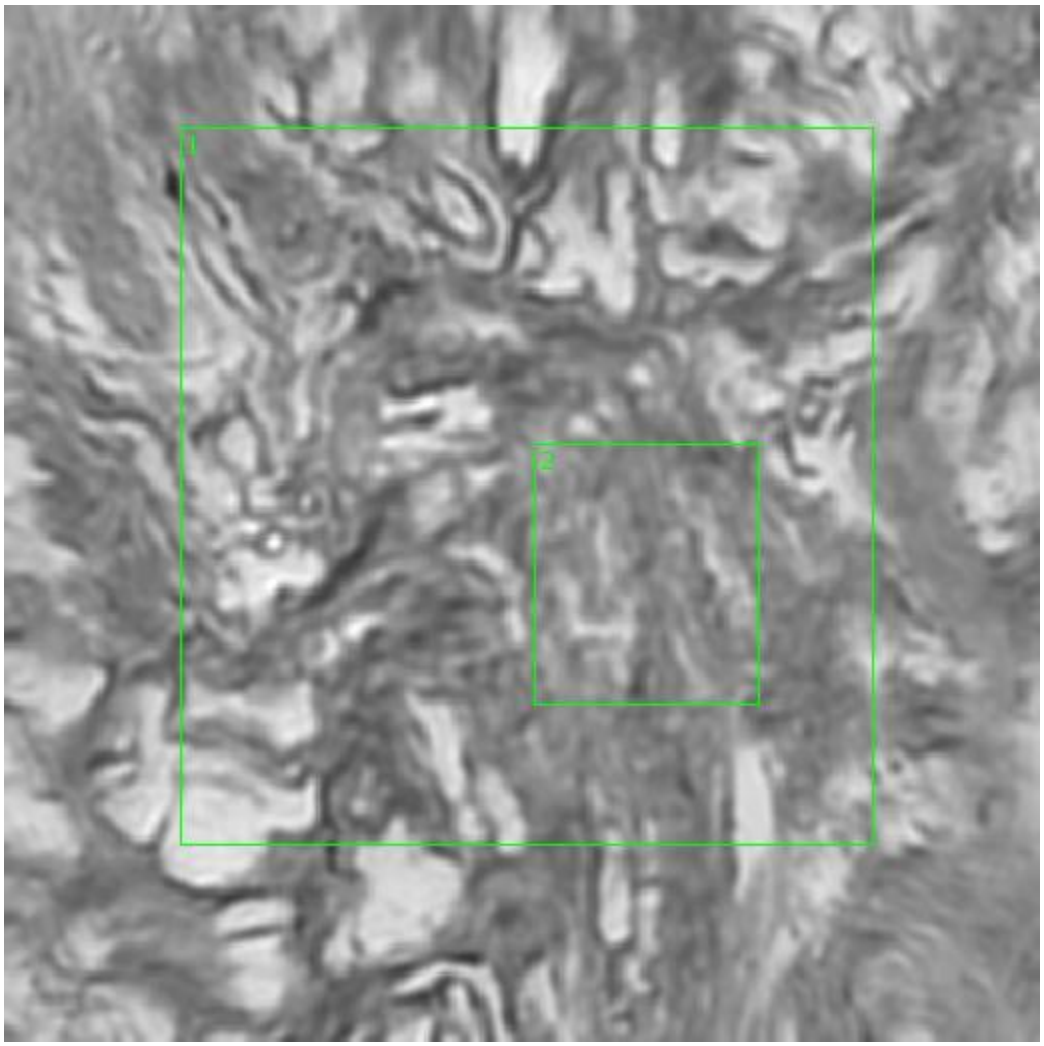



**Figure 0.13. Point K Regions of Interest**
**Region 1: normal right breast tissue**
**Region 2: concentrated region, normal right breast tissue**

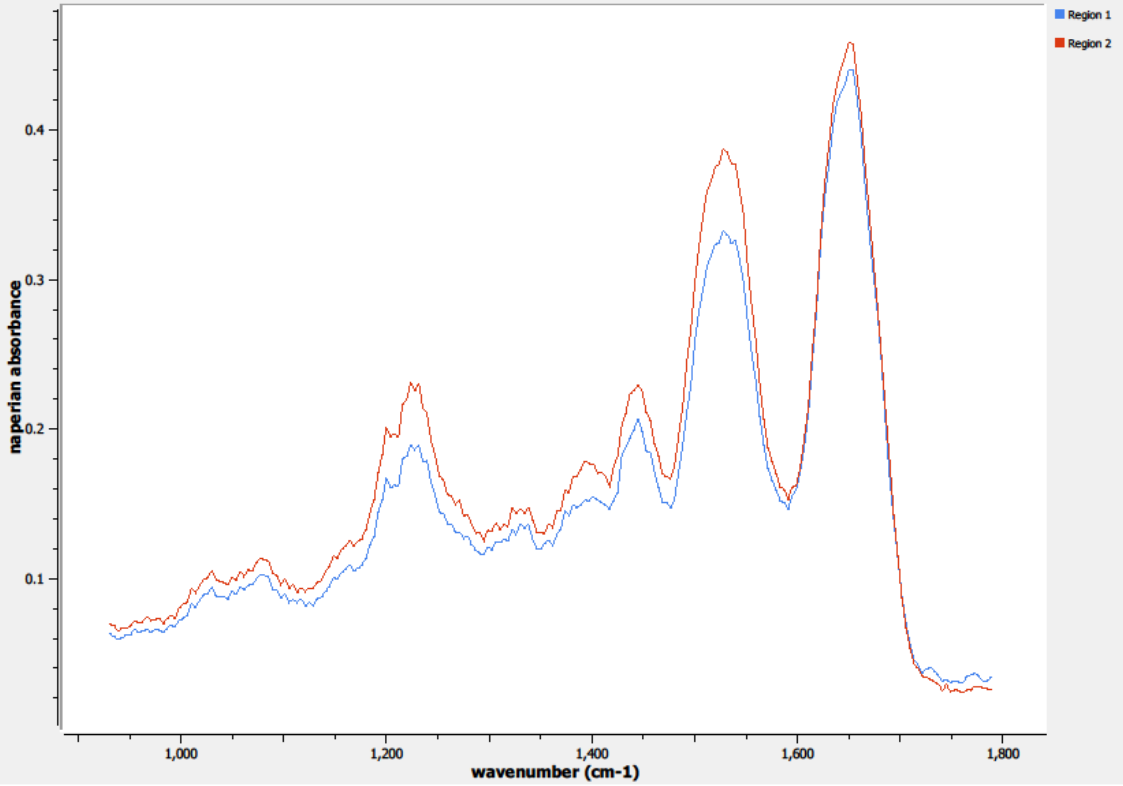



**Figure 0.14. Point K Spectra**

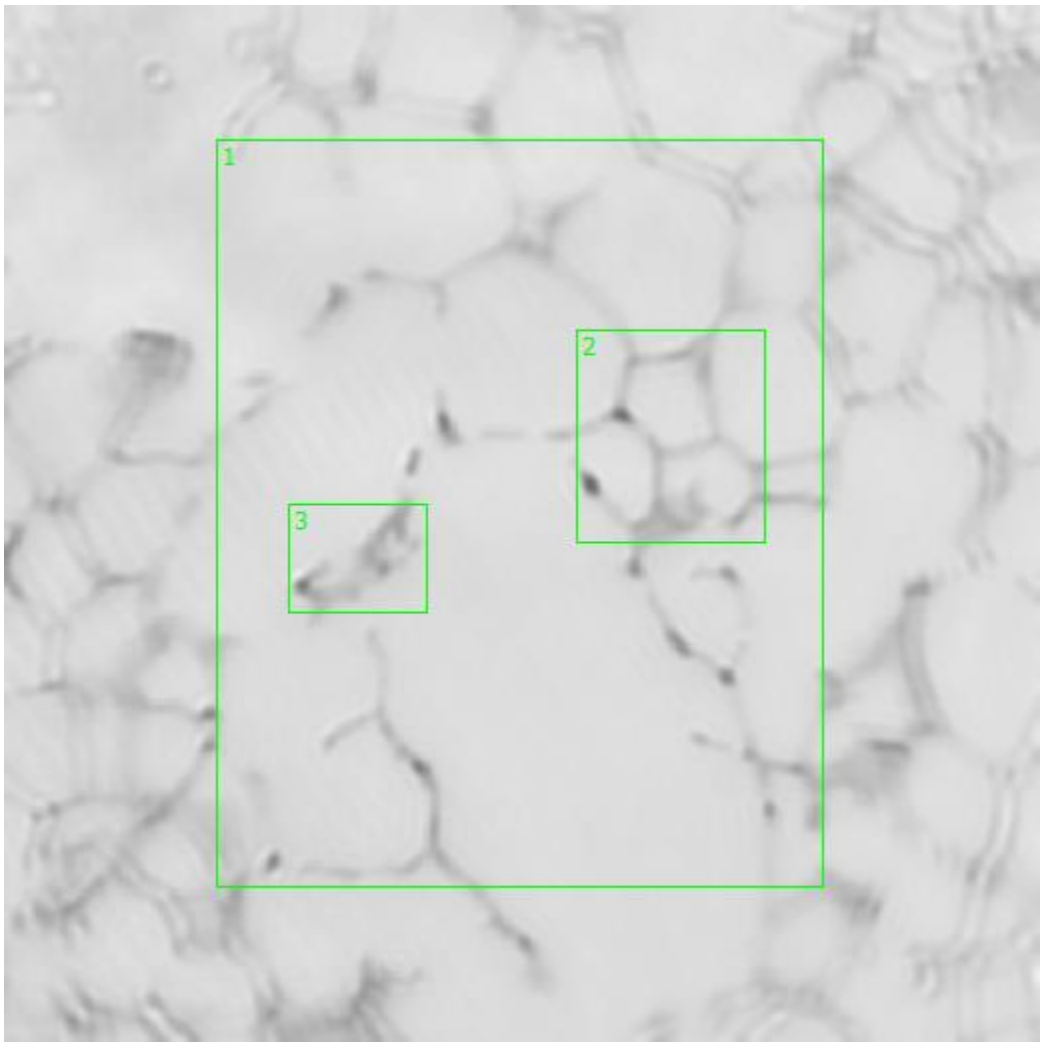



**Figure 0.15. Point L Regions of Interest**
**Region 1: normal right breast tissue, washed out fat cells**
**Region 2, 3: concentrated regions, normal right breast tissue, washed out fat cells**

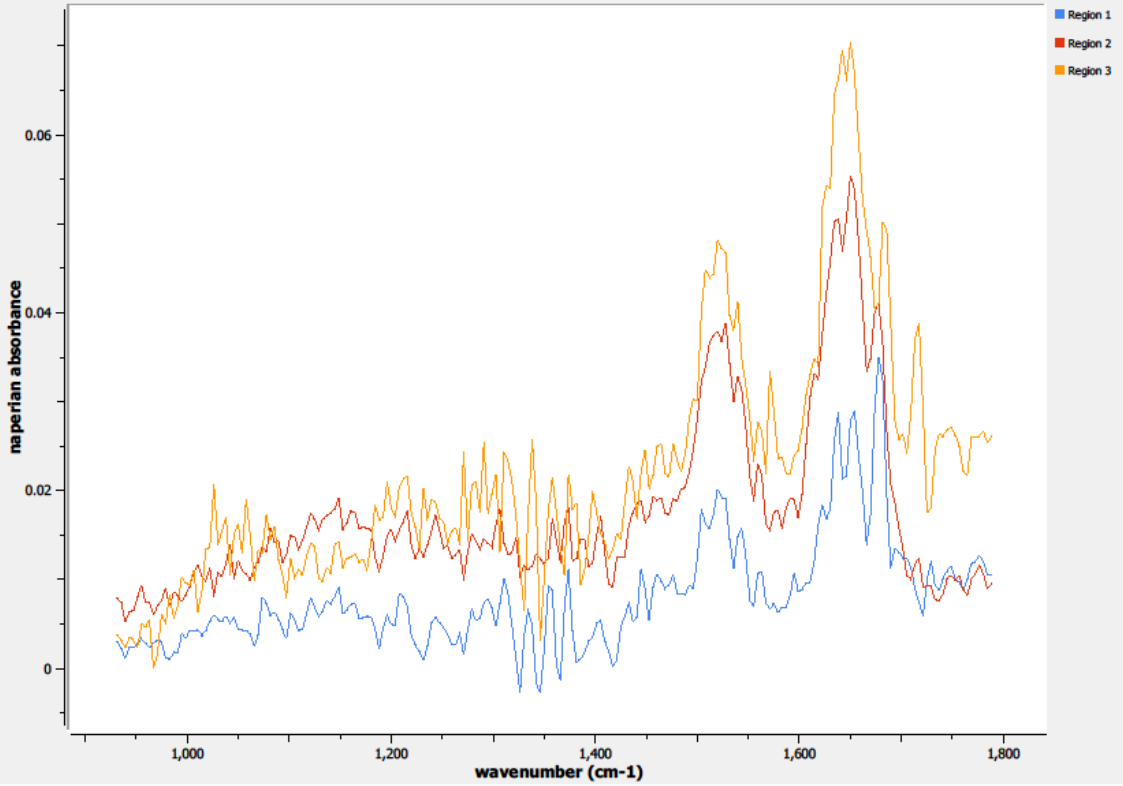



**Figure 0.16. Point L Spectra, washed out fat cells, unused in data results**

## A3.12-03-A140d

| | |
|---|---|
| Unique ID | 12-03-A140d |
| Age | 58 |
| Race | white |
| Location of Sample Collection | right breast |
| Notes of Interest | none |

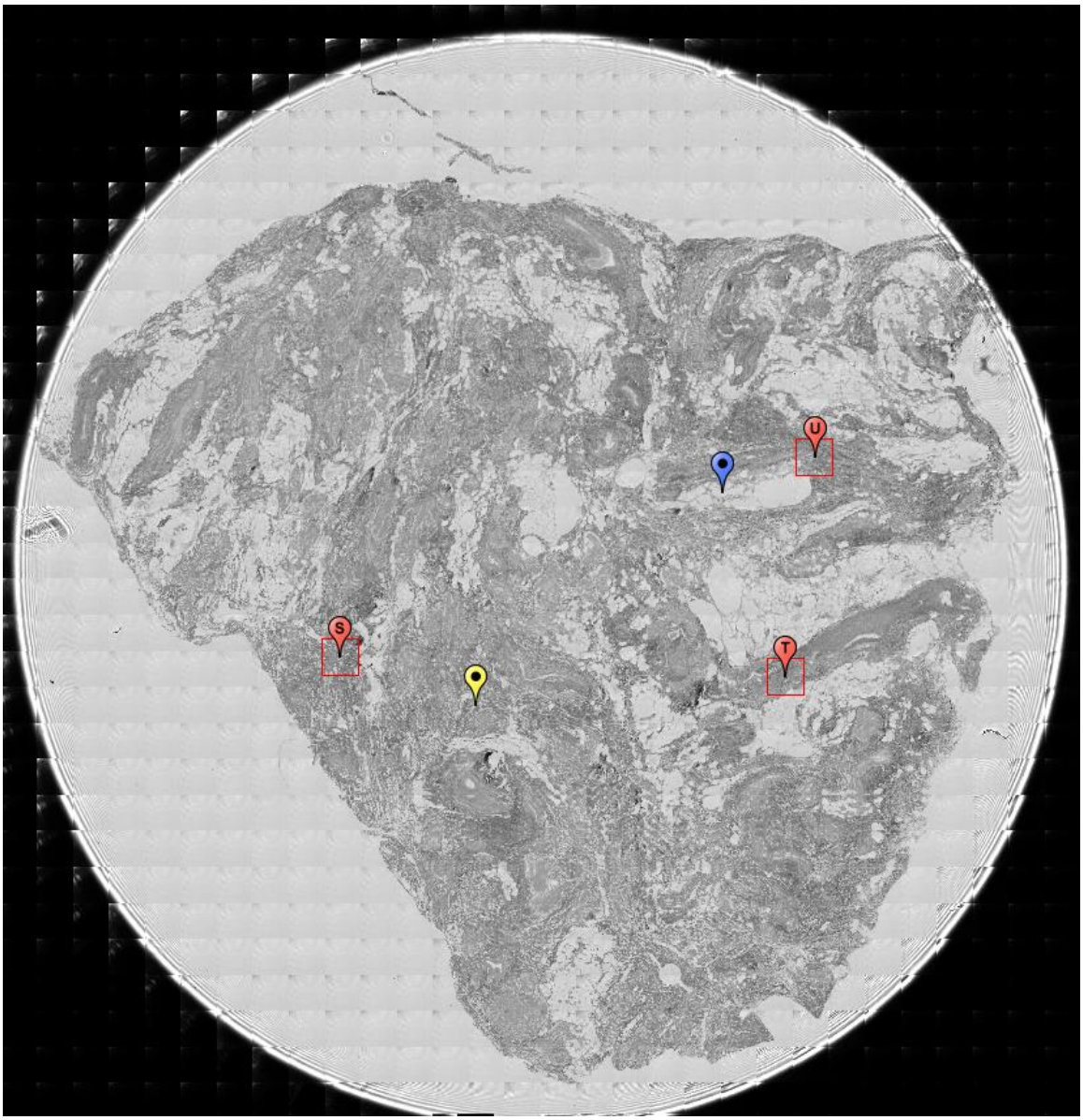



**Figure 0.17. Specimen 12-03-A140d, normal right breast tissue.**

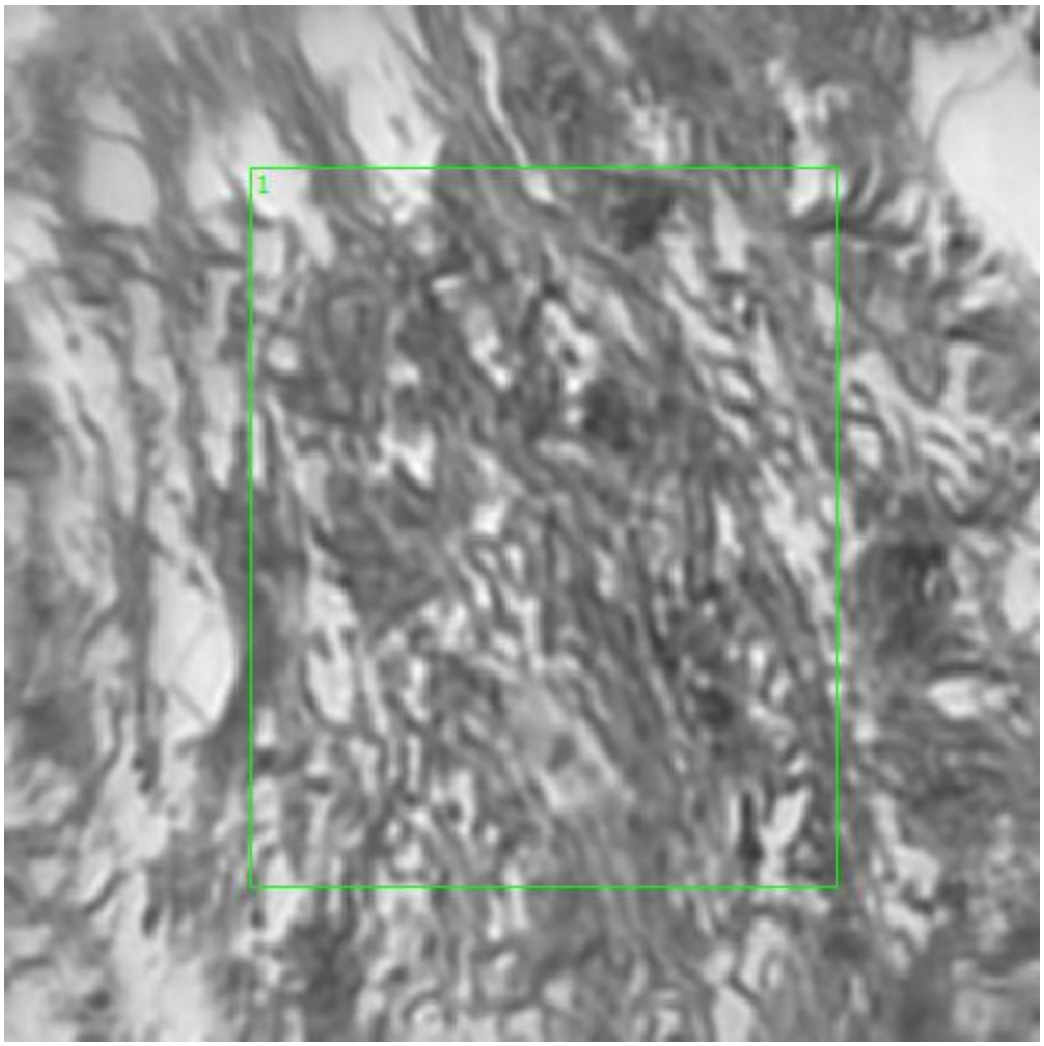



**Figure 0.18. Point S Regions of Interest**
**Region 1: right breast**

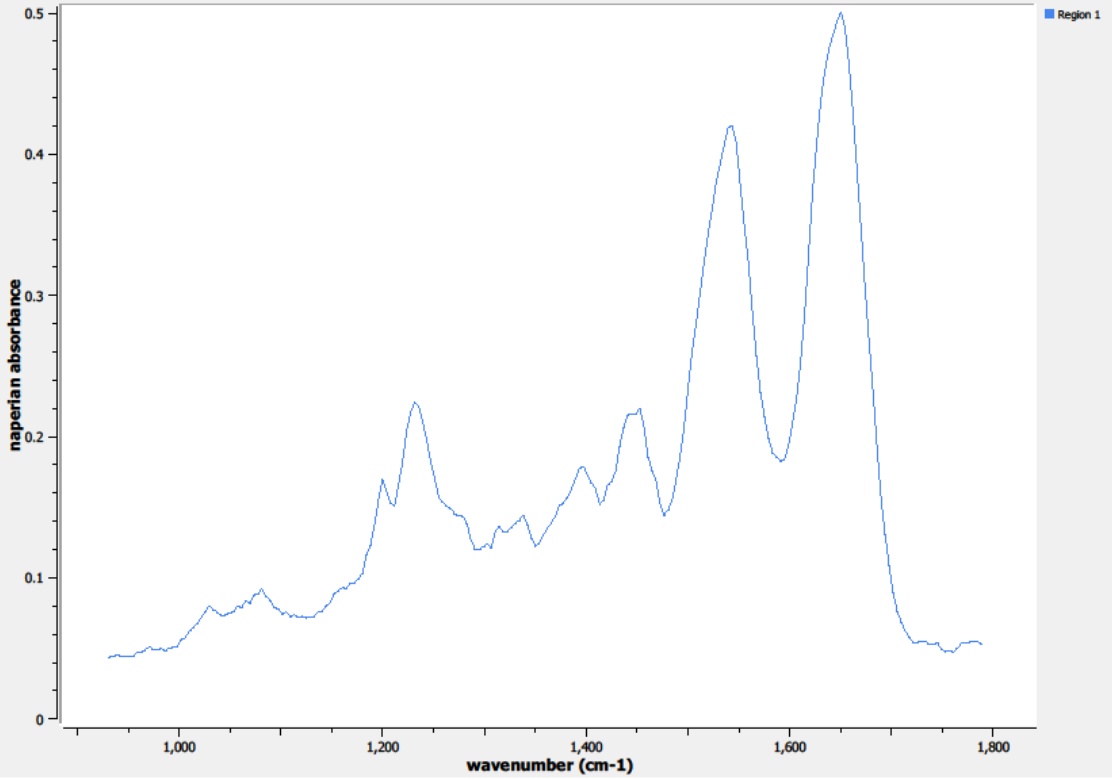



**Figure 0.19. Point S Spectra**

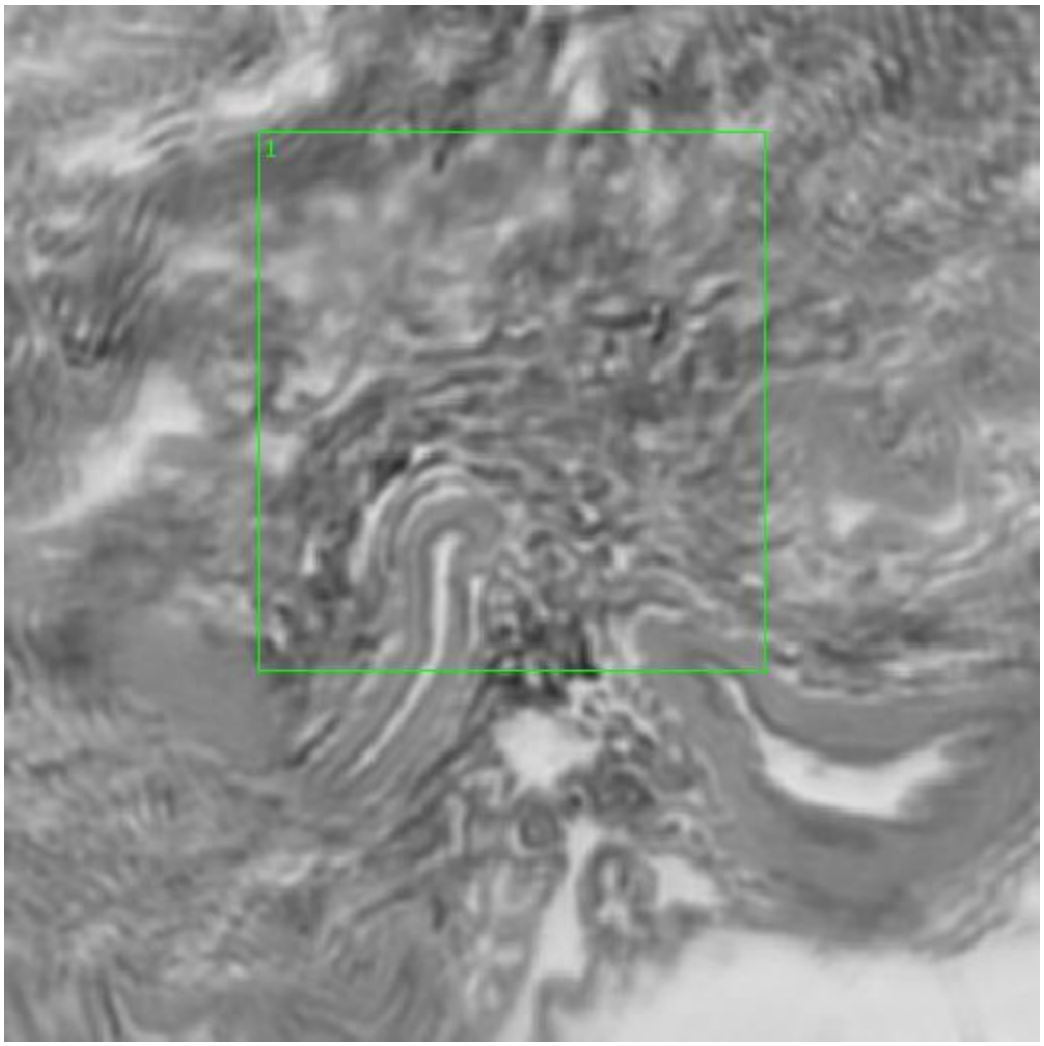



**Figure 0.20. Point T Regions of Interest**
**Region 1: right breast**

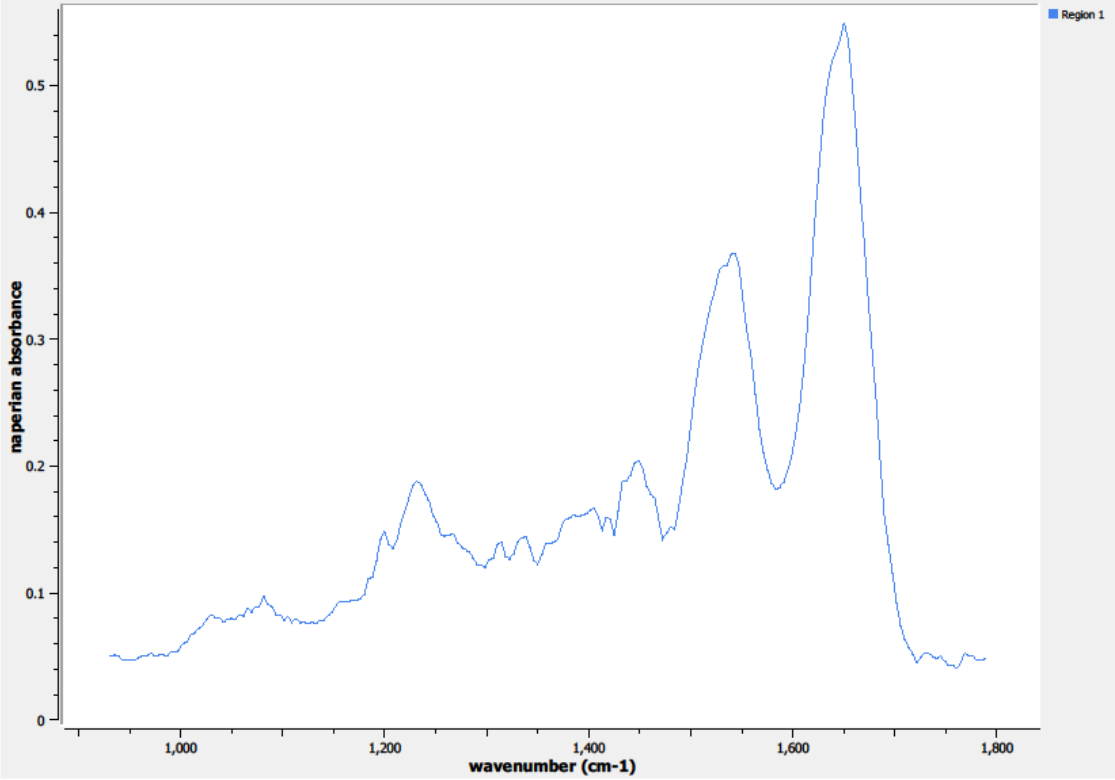



**Figure 0.21. Point T Spectra**

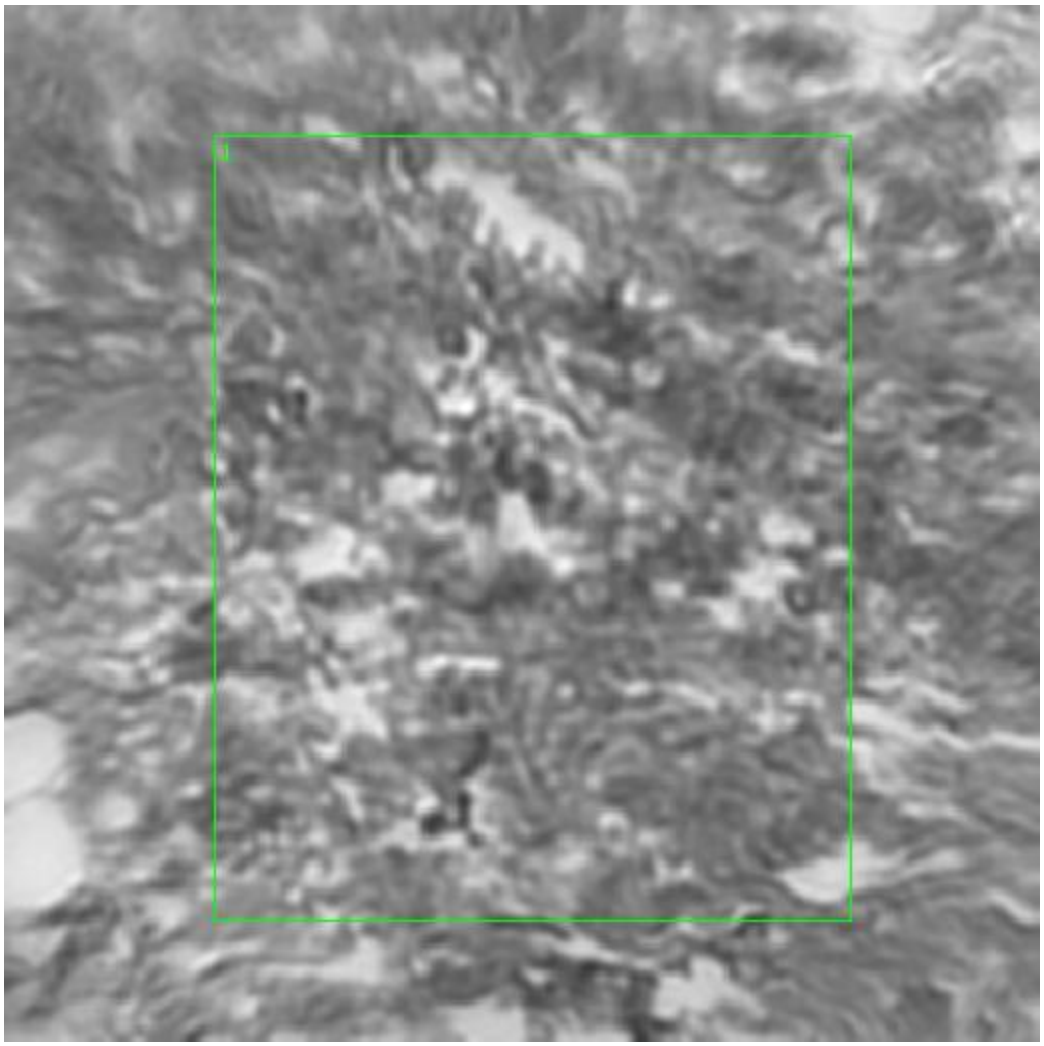


**Figure 0.22. Point U Regions of Interest**
**Region 1: right breast**

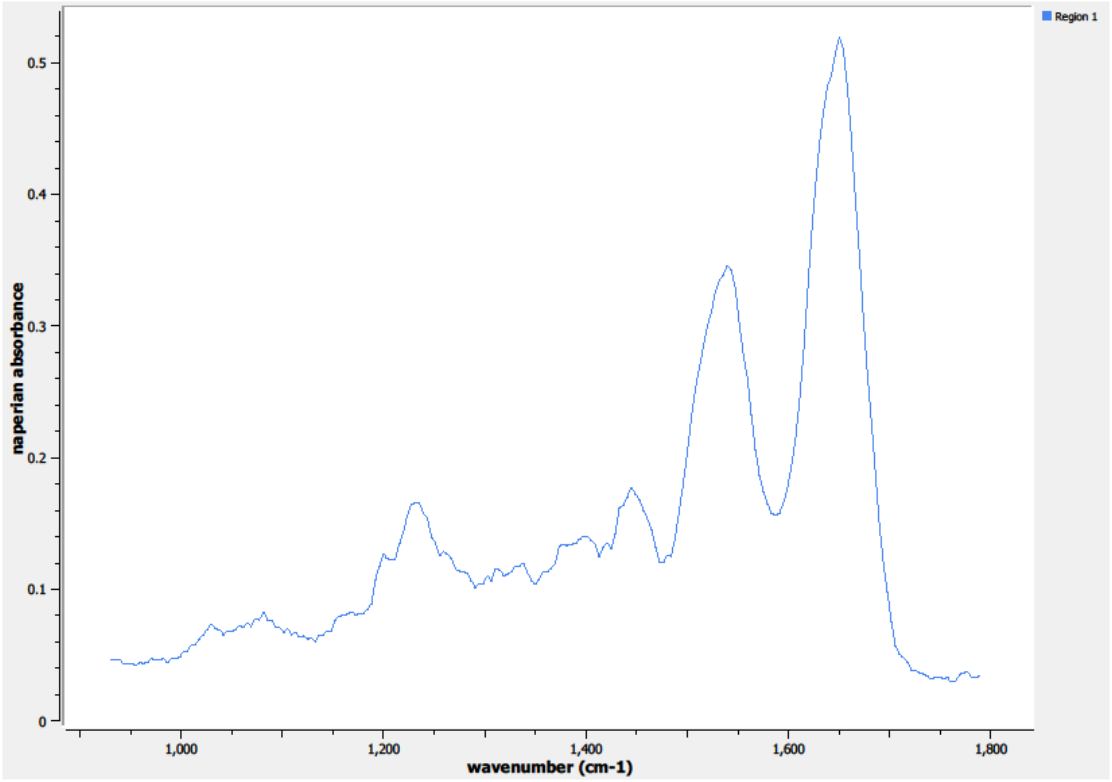



**Figure 0.23. Point U Spectra**

## A4.12-04-A020d

| | |
|---|---|
| Unique ID | 12-04-A020d |
| Age | 60 |
| Race | black |
| Location of Sample Collection | right breast |
| Notes of Interest | none |

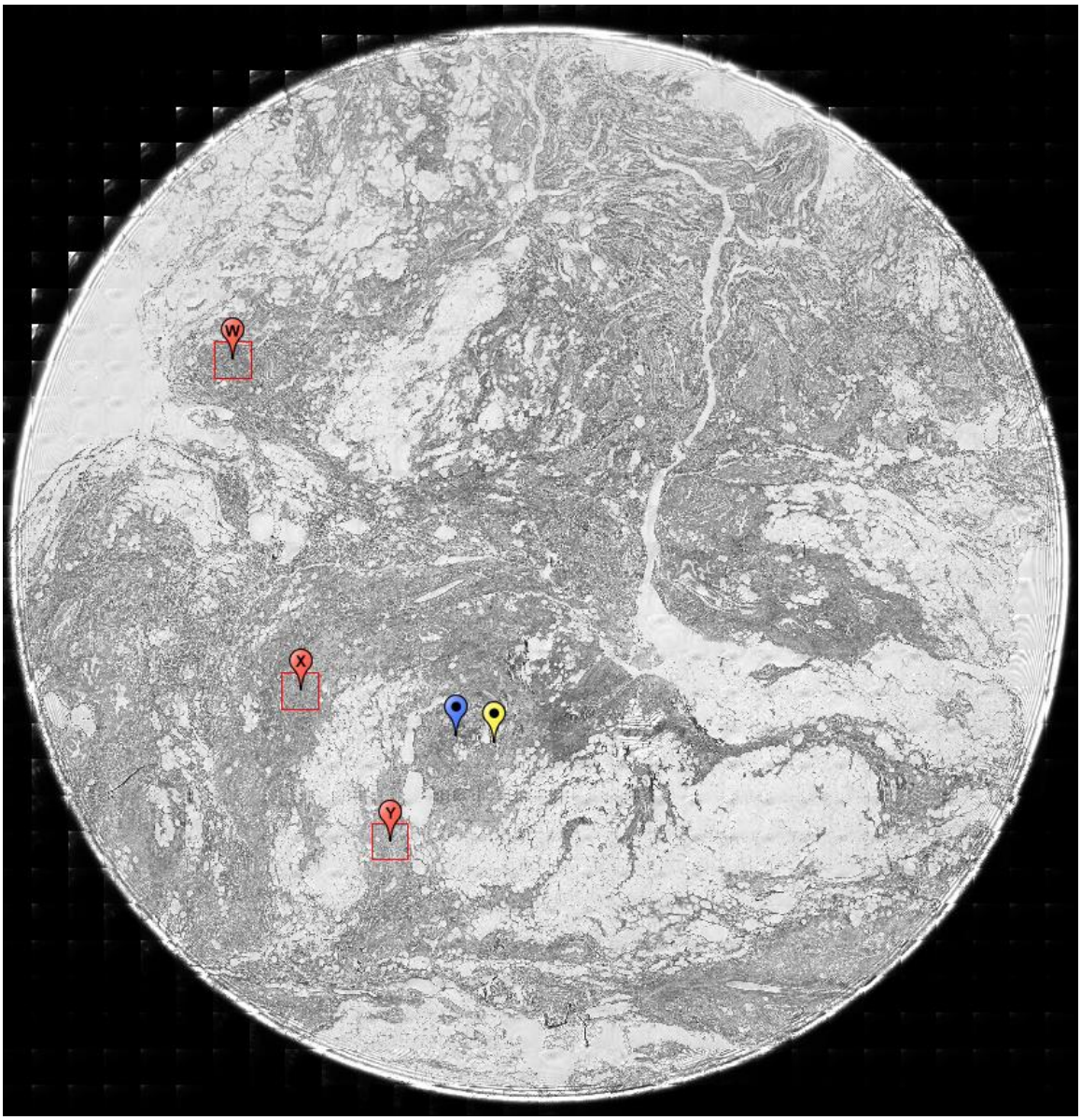



**Figure 0.24. Specimen 12-04-A020d, normal right breast tissue.**

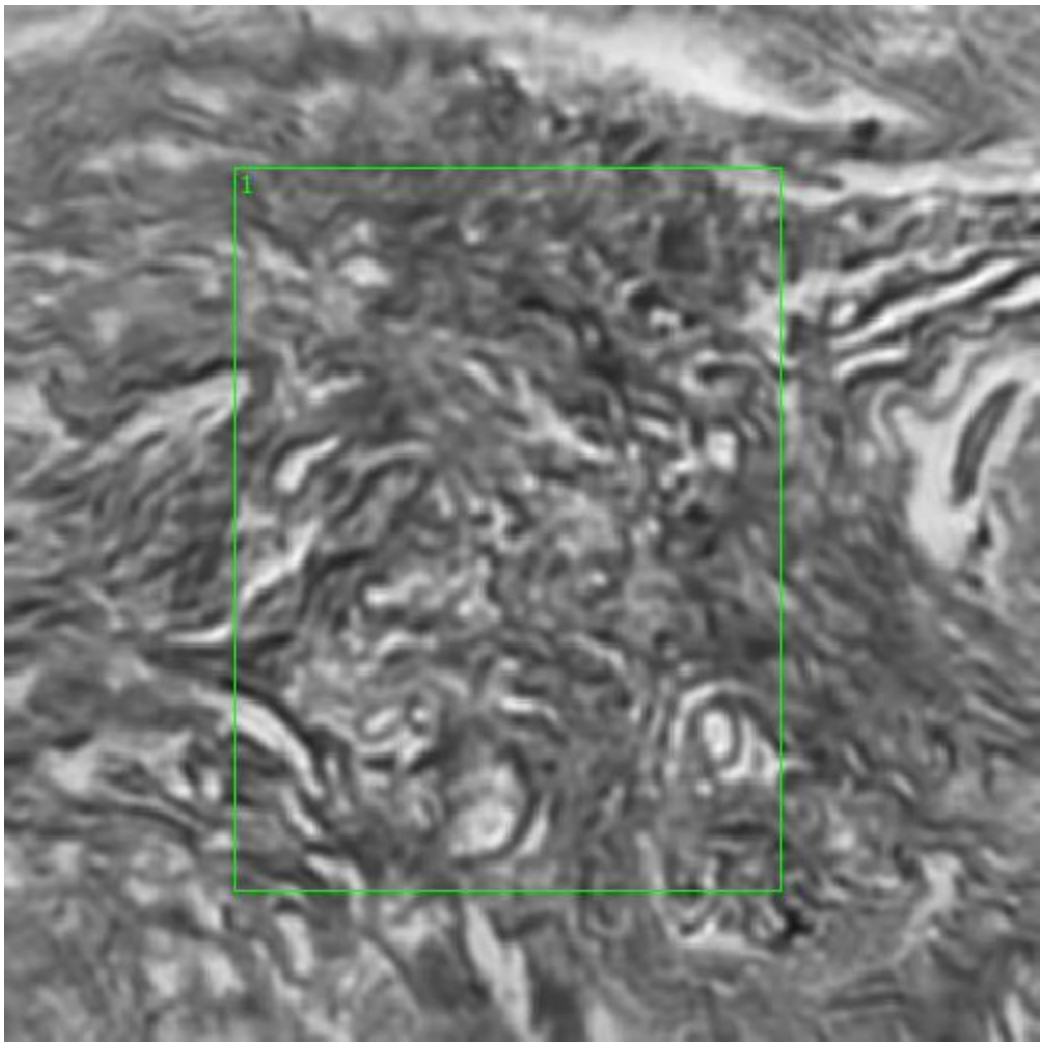



**Figure 0.25. Point W Regions of Interest**
**Region 1: right breast**

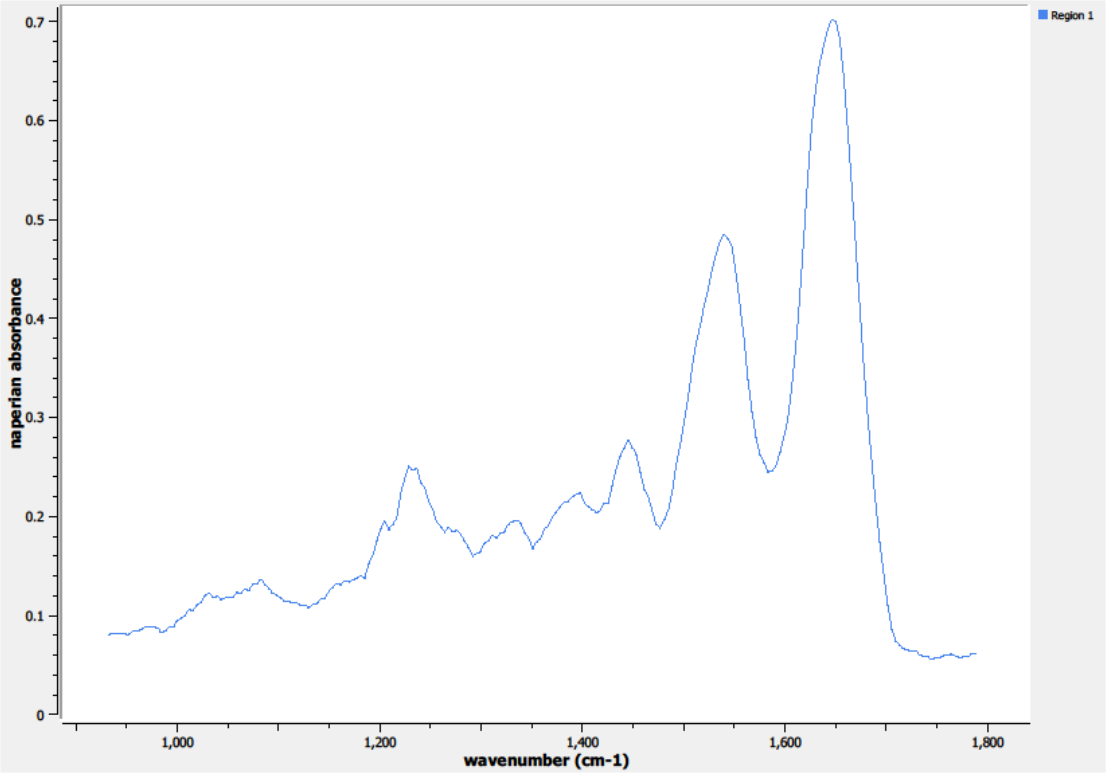



**Figure 0.26. Point W Spectra**

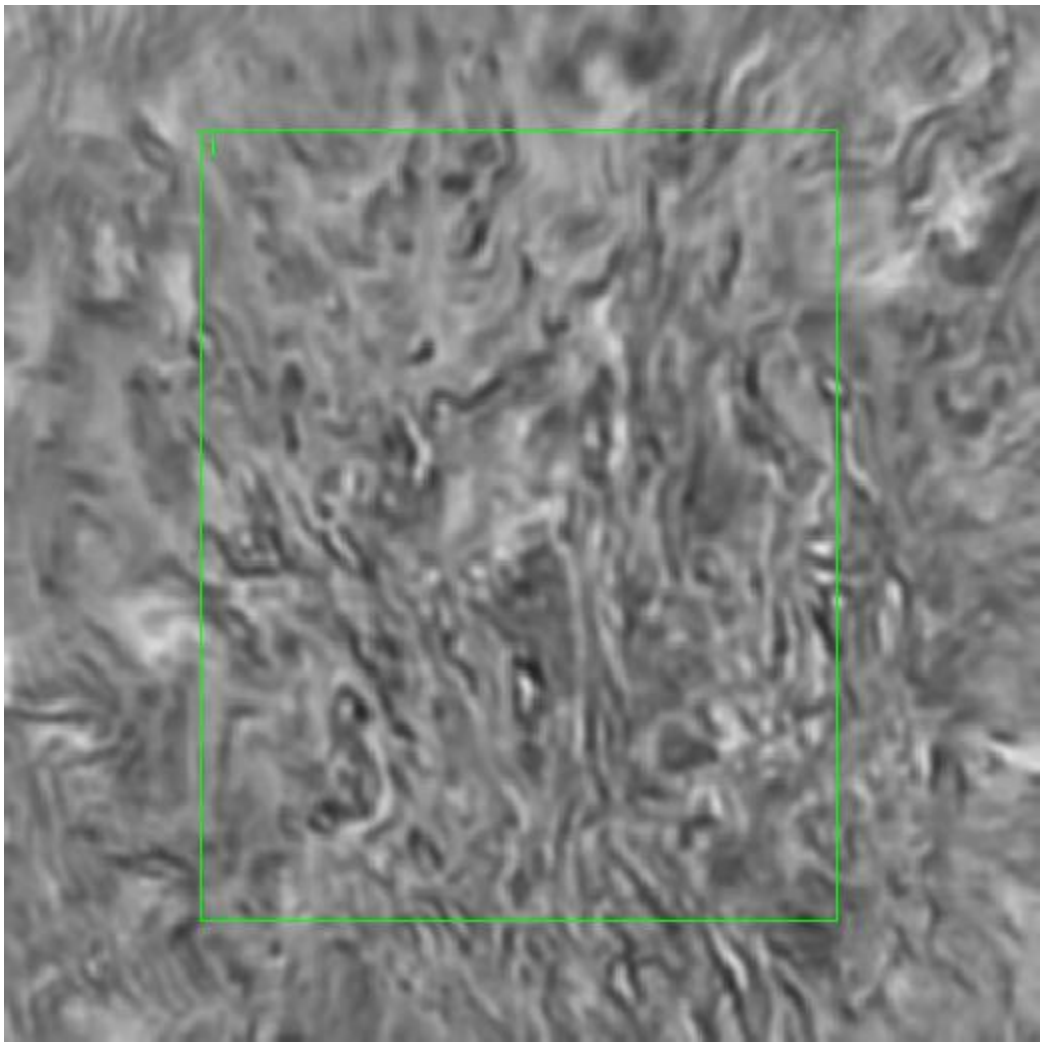


**Figure 0.27. Point X Regions of Interest**
**Region 1: right breast**

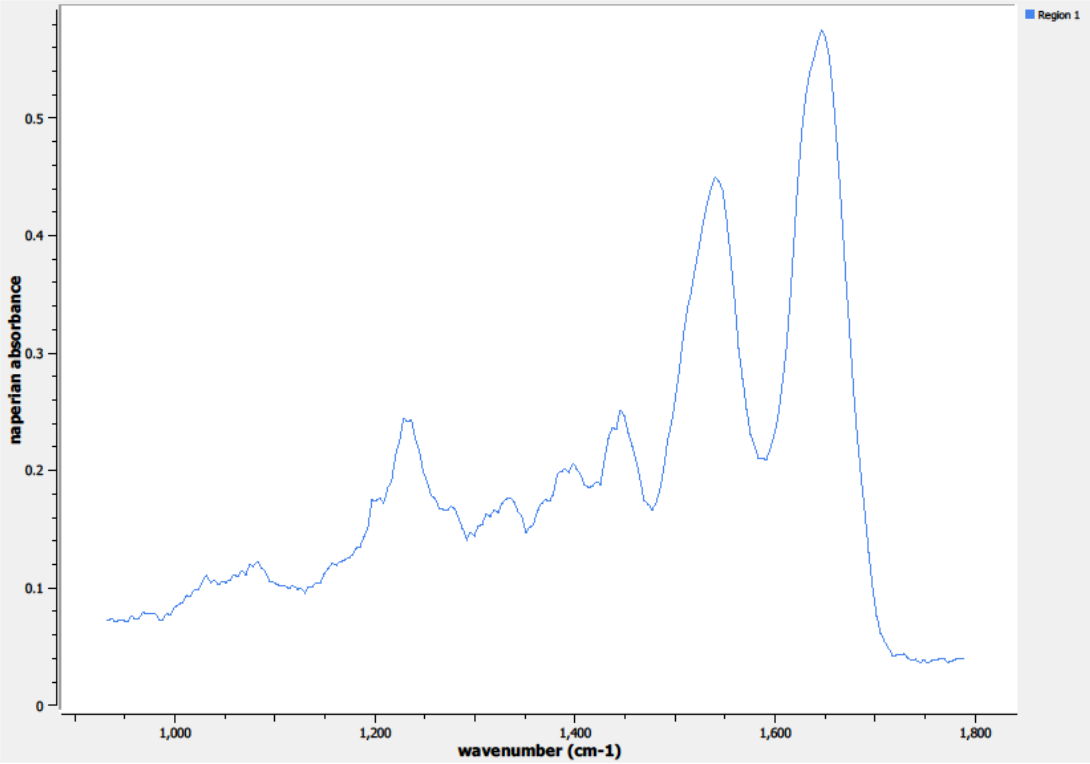



**Figure 0.28. Point X Spectra**

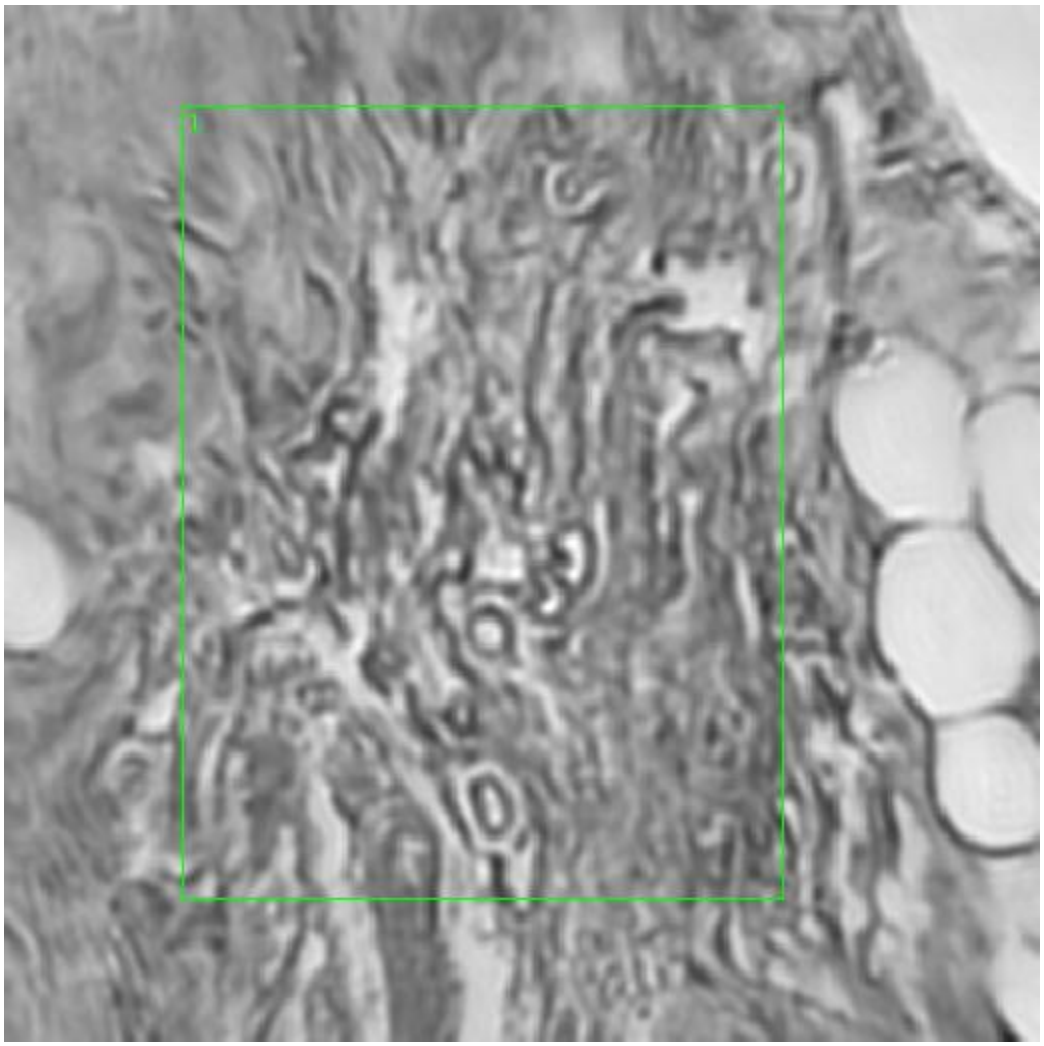


**Figure 0.29. Point Y Regions of Interest**
**Region 1: right breast**

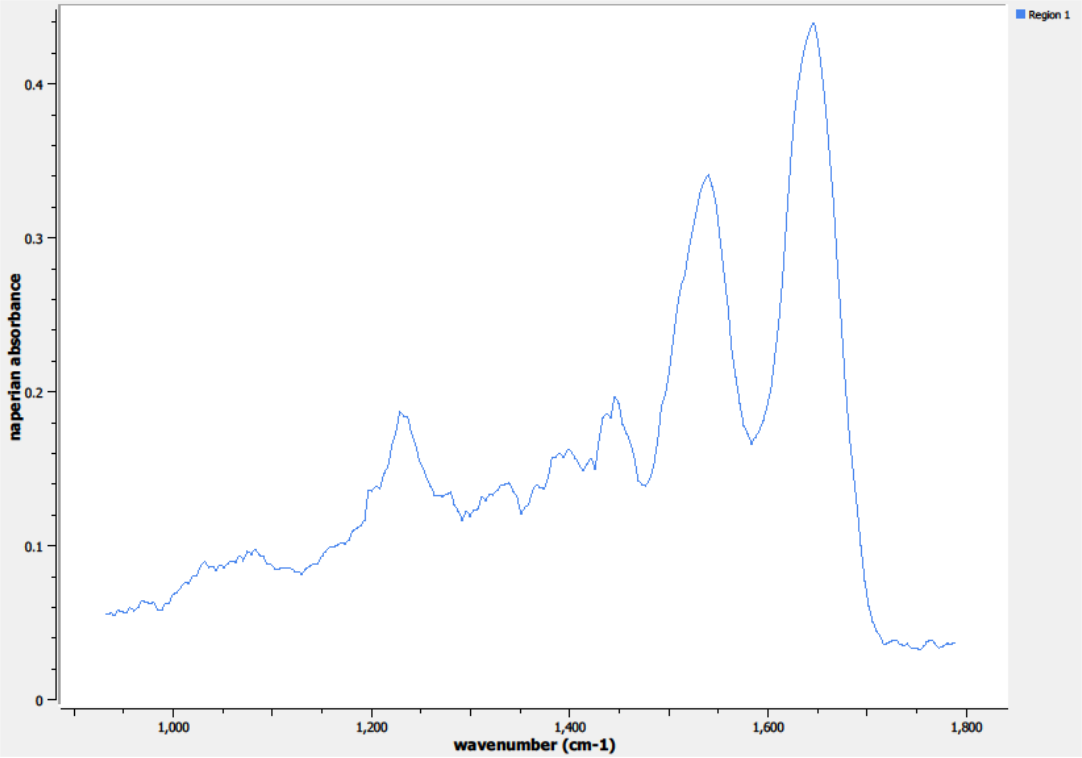



**Figure 0.30. Point Y Spectra**

## A5.12-05-A058d

| | |
|---|---|
| Unique ID | 12-05-A058d |
| Age | 62 |
| Race | white |
| Location of Sample Collection | left breast |
| Notes of Interest | abnormal tissue sample, may not use |

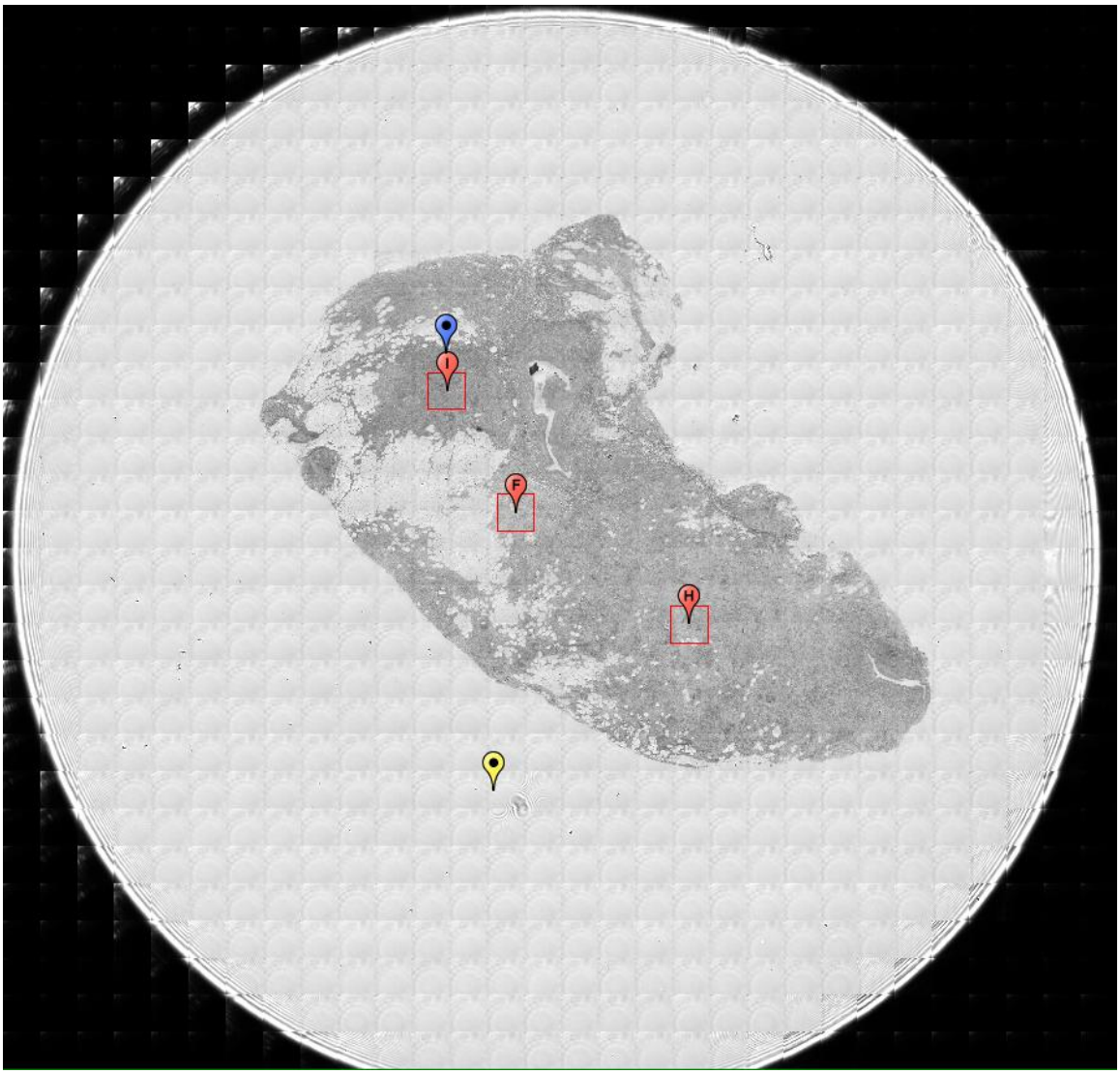



**Figure 0.31. Specimen 12-05-A058d, normal left breast tissue.**

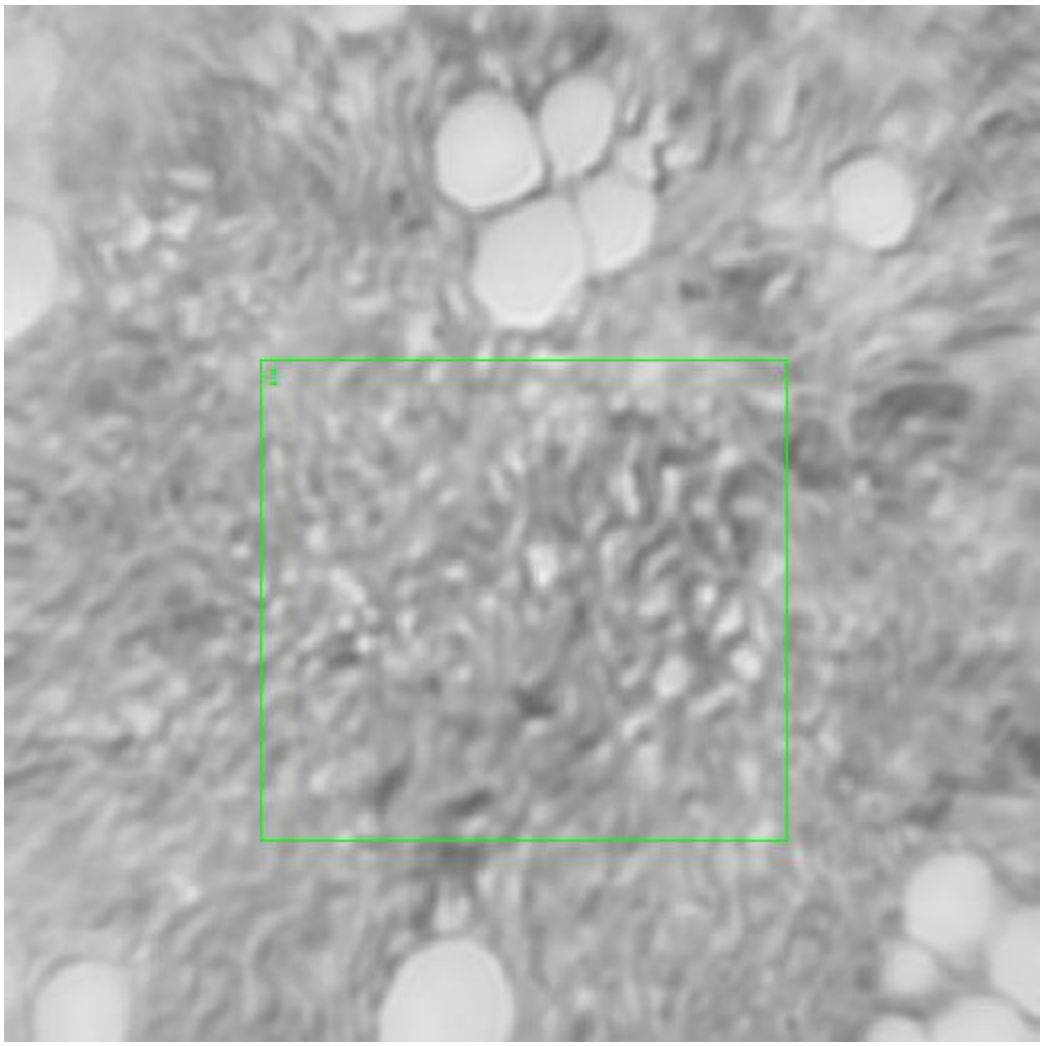



**Figure 0.32. Point F Regions of Interest**
**Region 1: left breast**

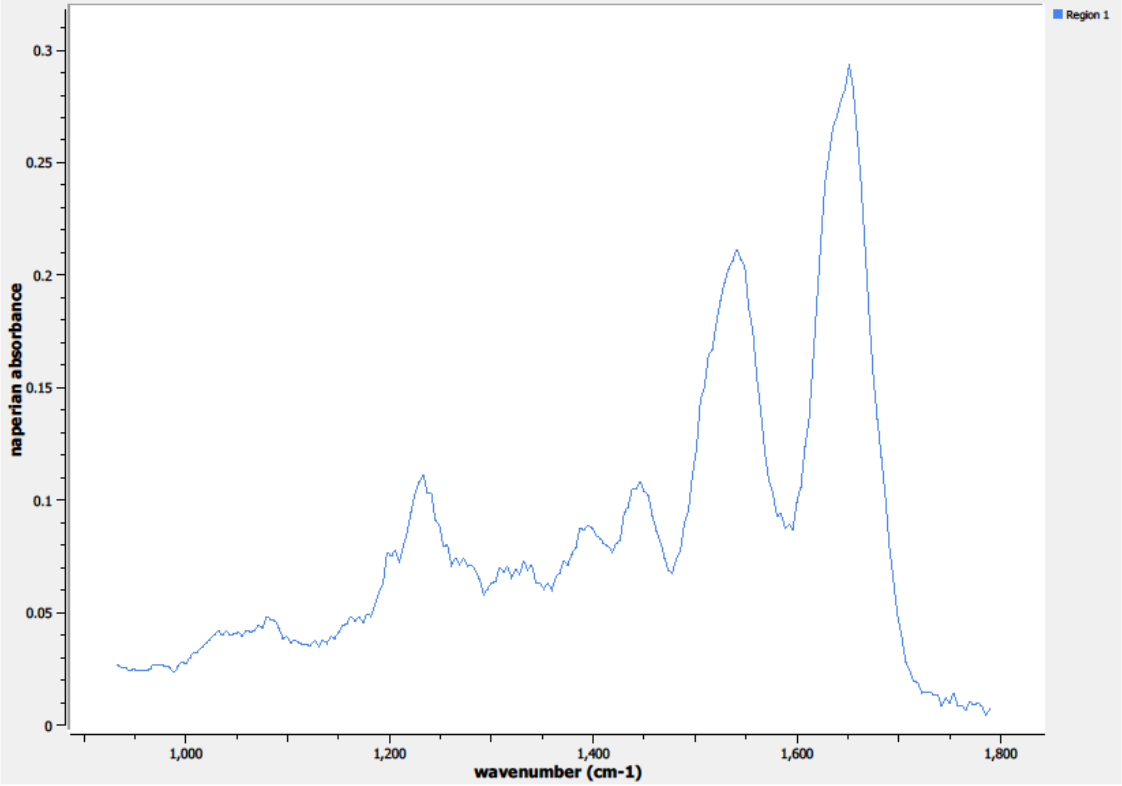



**Figure 0.33. Point F Spectra**

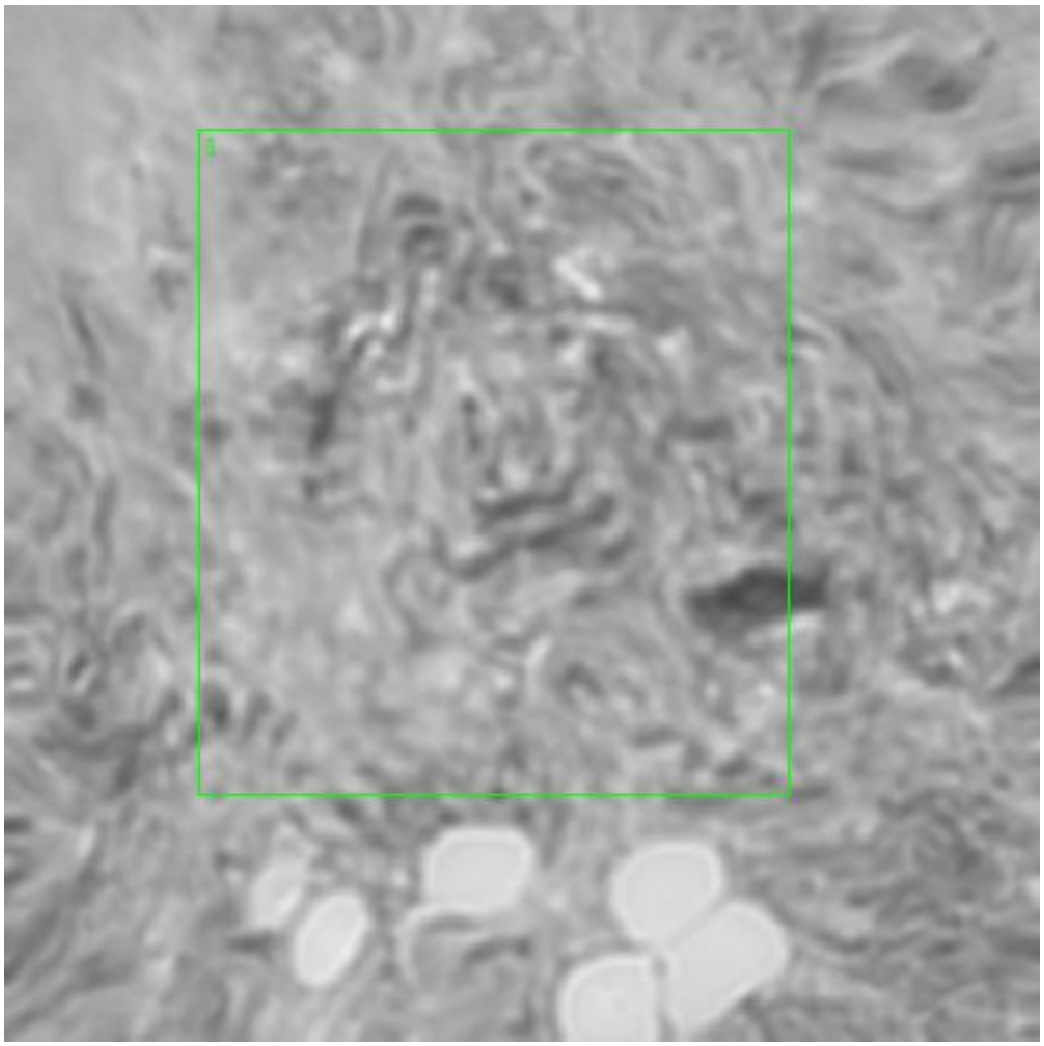


**Figure 0.34. Point H Regions of Interest**
**Region 1: left breast**

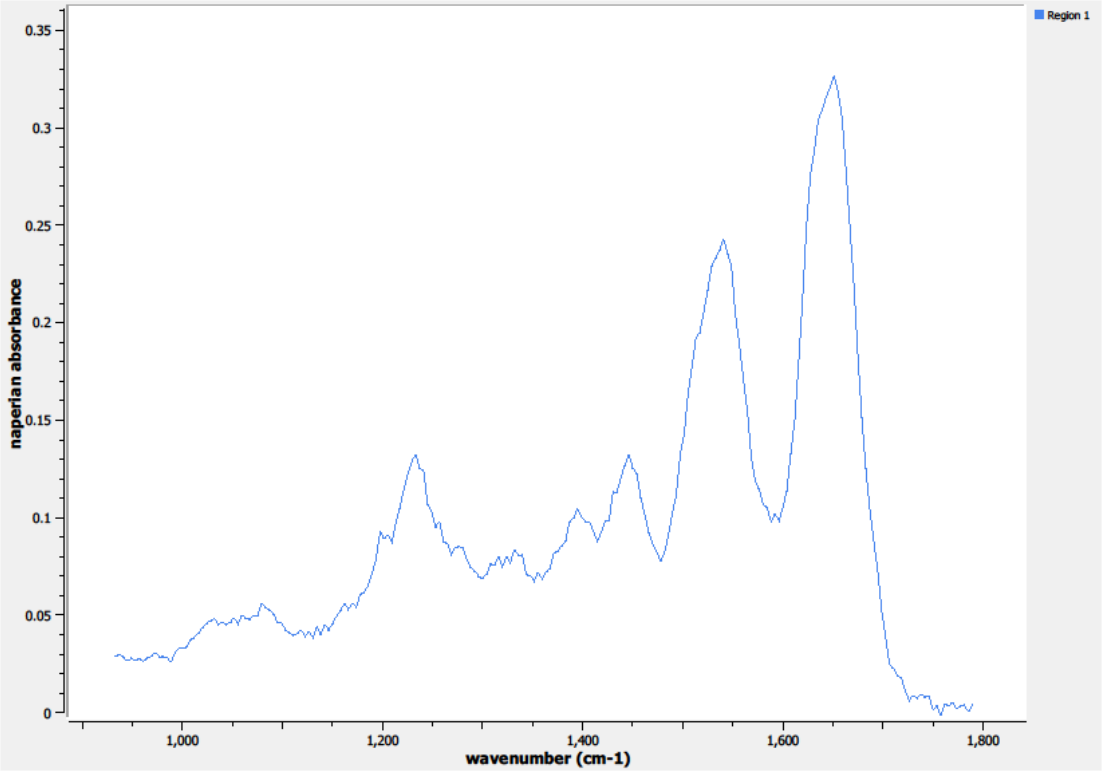



**Figure 0.35. Point H Spectra**

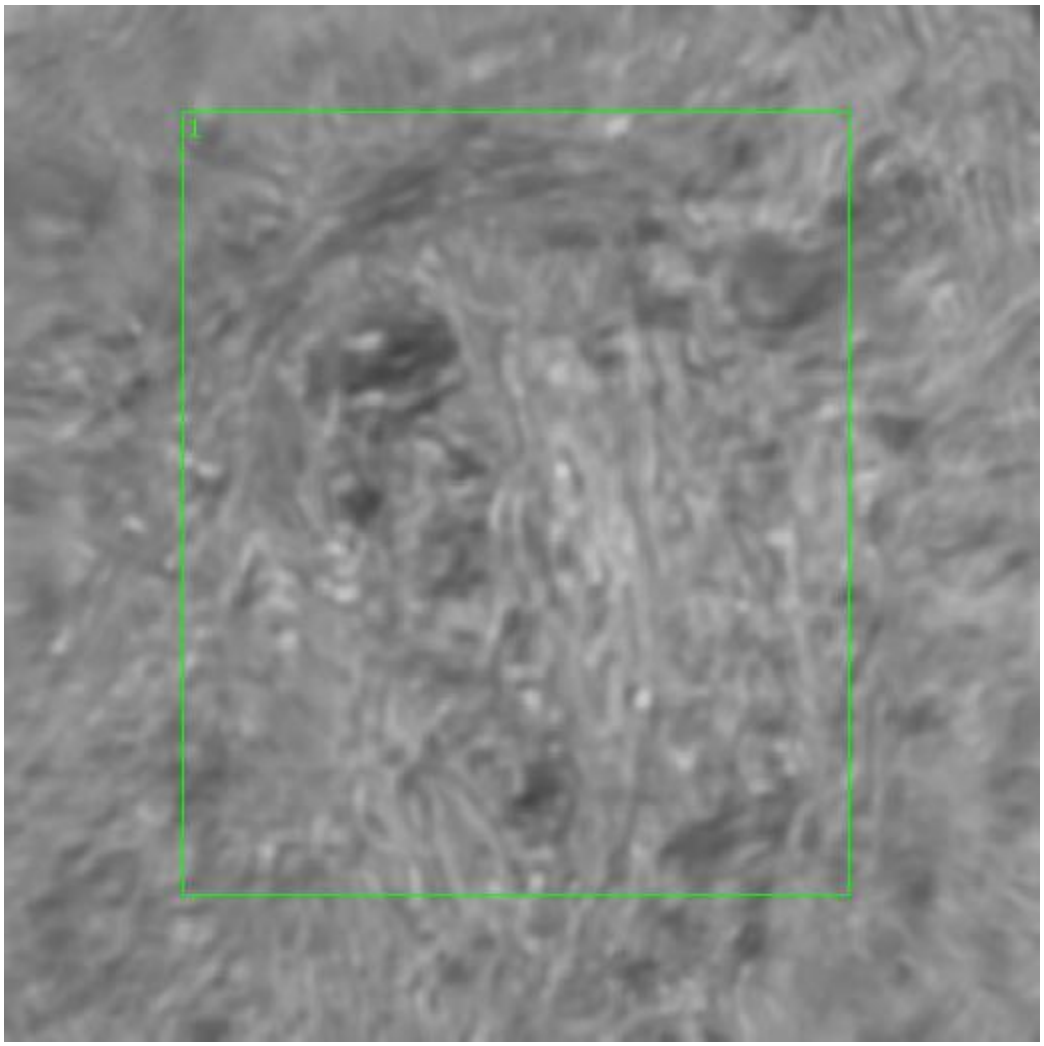



**Figure 0.36. Point I Regions of Interest**
**Region 1: left breast**

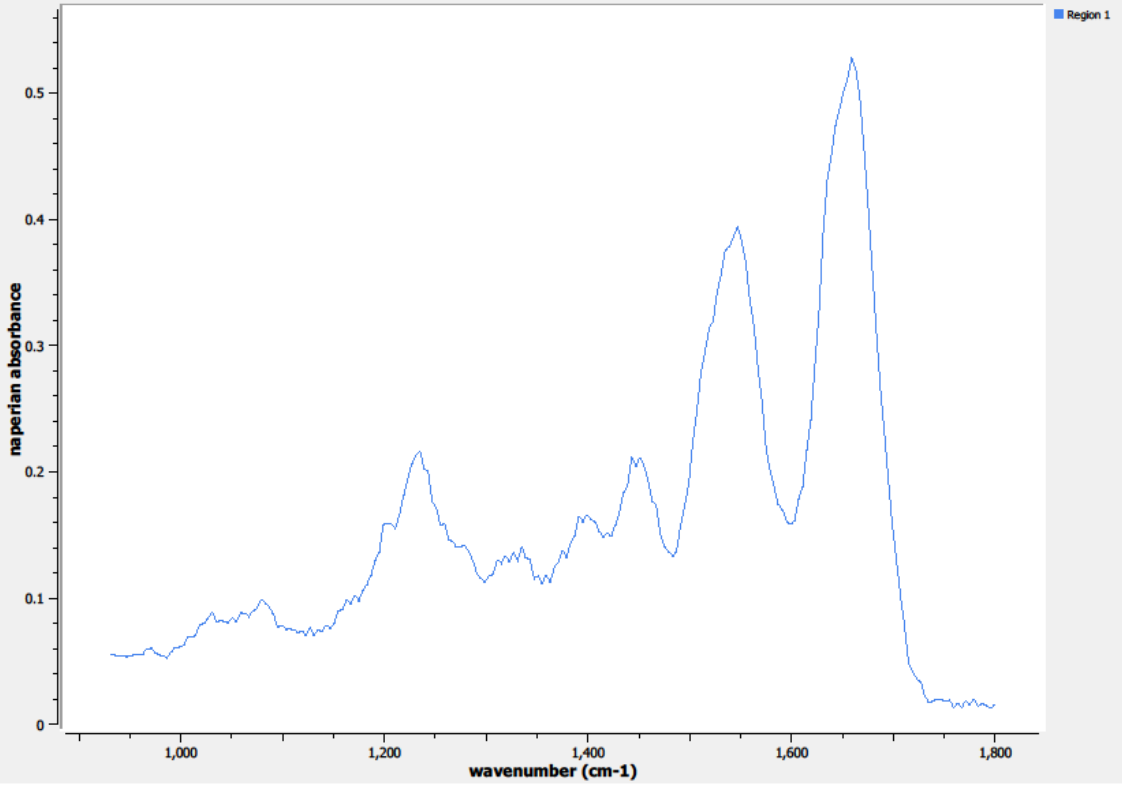



**Figure 0.37. Point I Spectra**

## A6.12-10-A071a

| | |
|---|---|
| Unique ID | 12-10-A071a |
| Age | 64 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

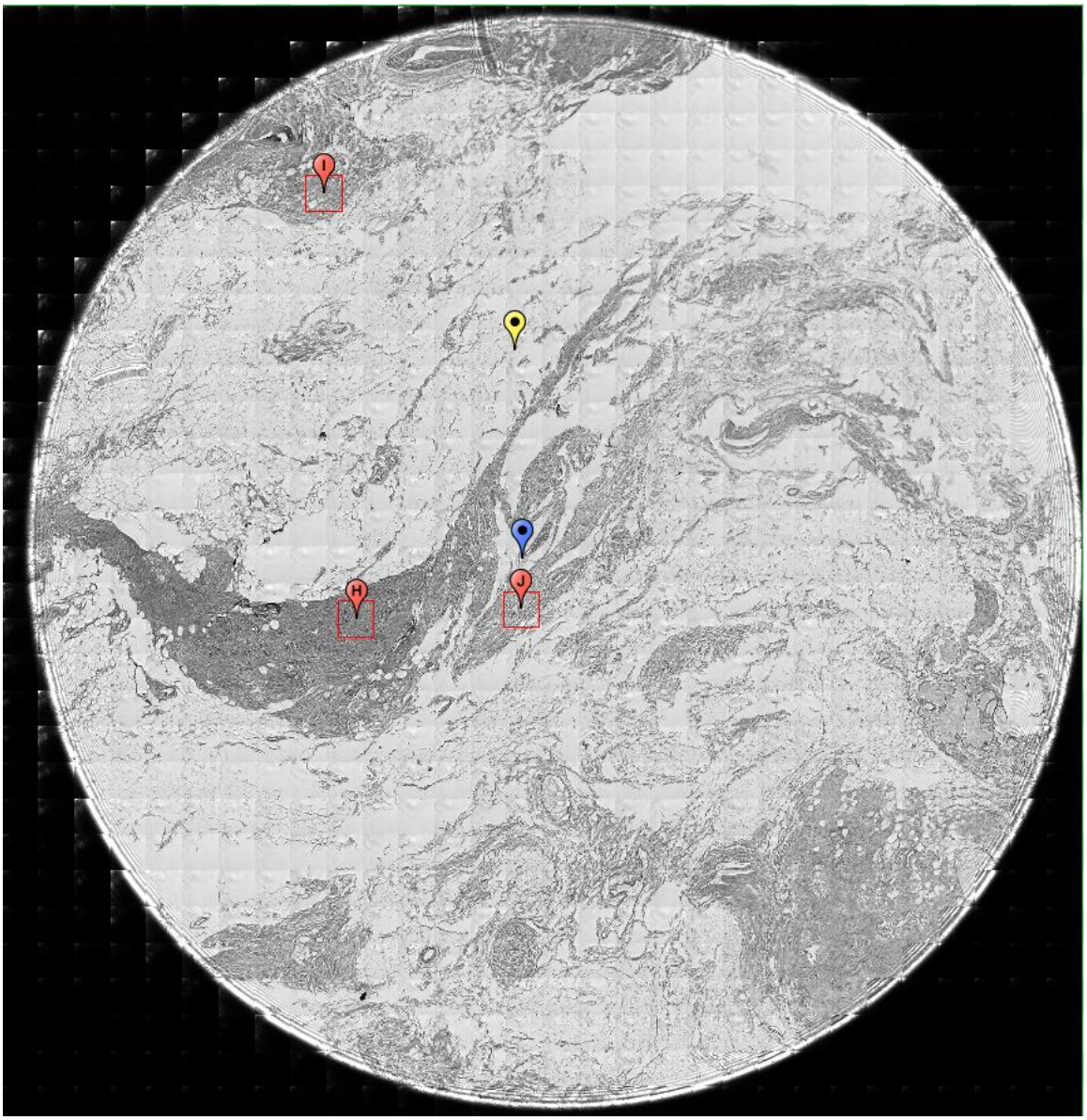



**Figure 0.38. Specimen 12-10-A071a, normal breast tissue.**

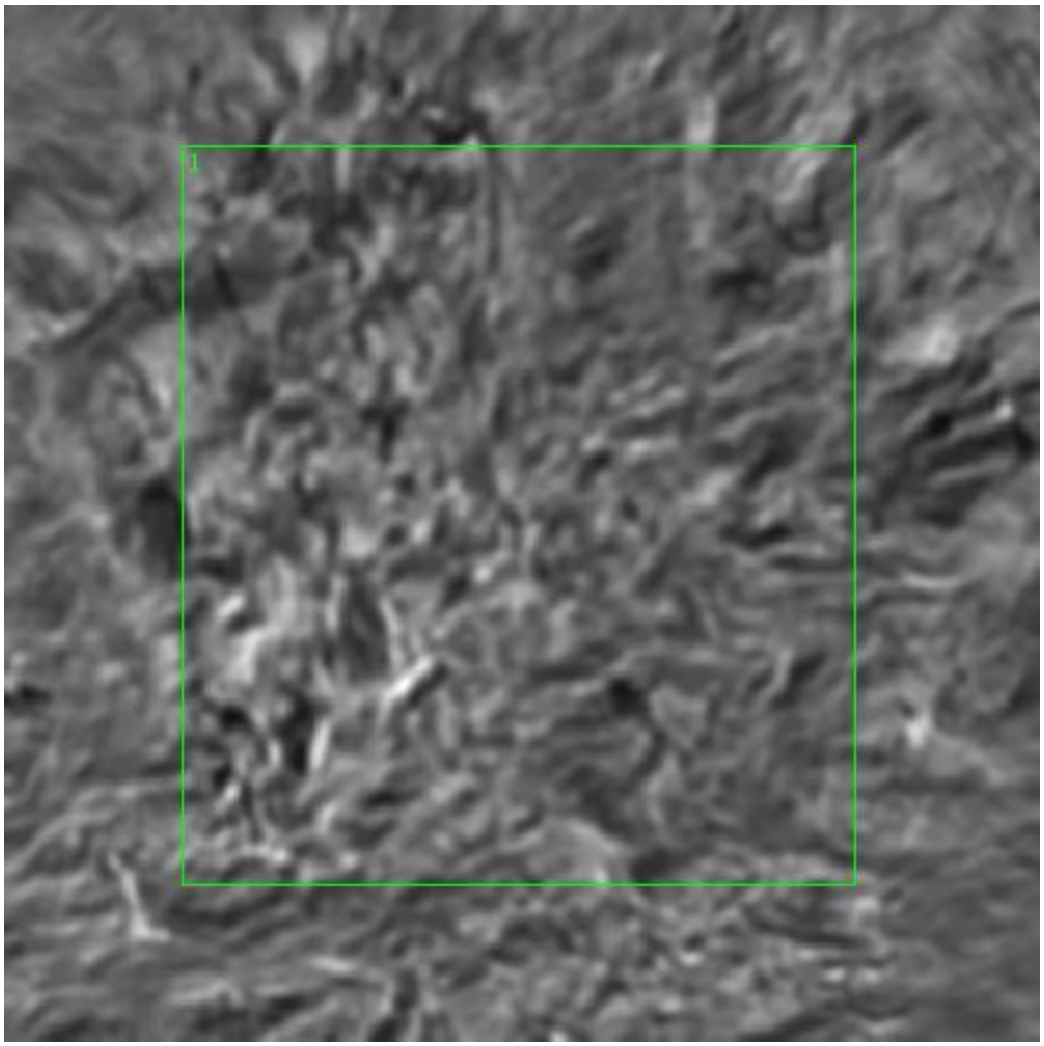


**Figure 0.39. Point H Regions of Interest**
**Region 1: breast**

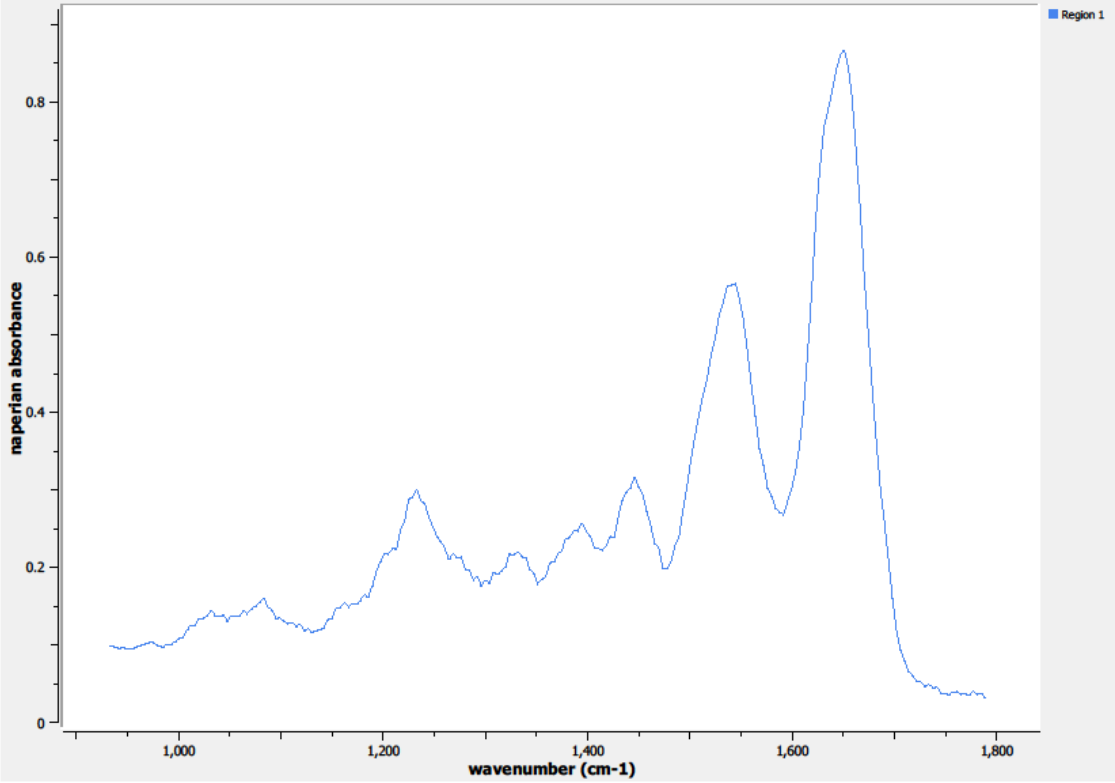



**Figure 0.40. Point H Spectra**

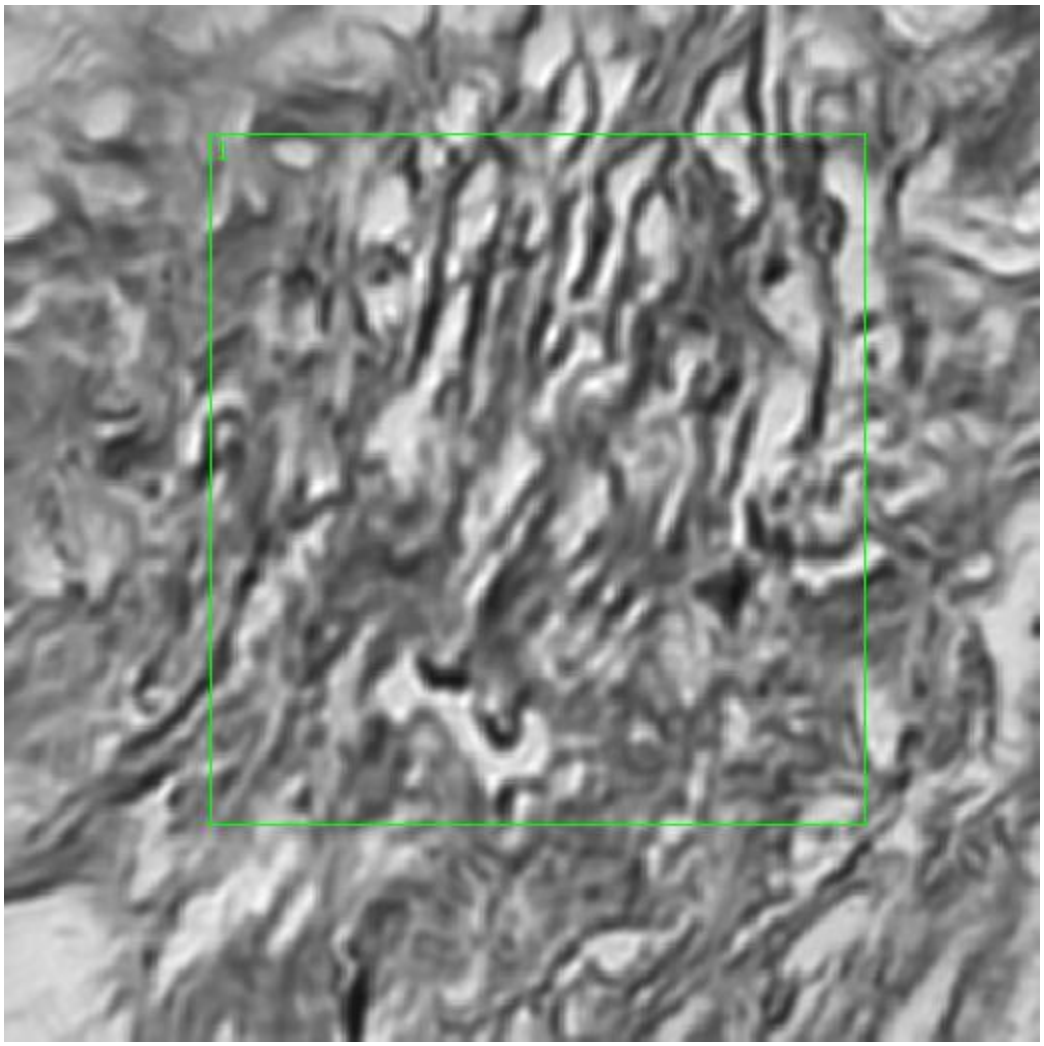


**Figure 0.41. Point I Regions of Interest**
**Region 1: breast**

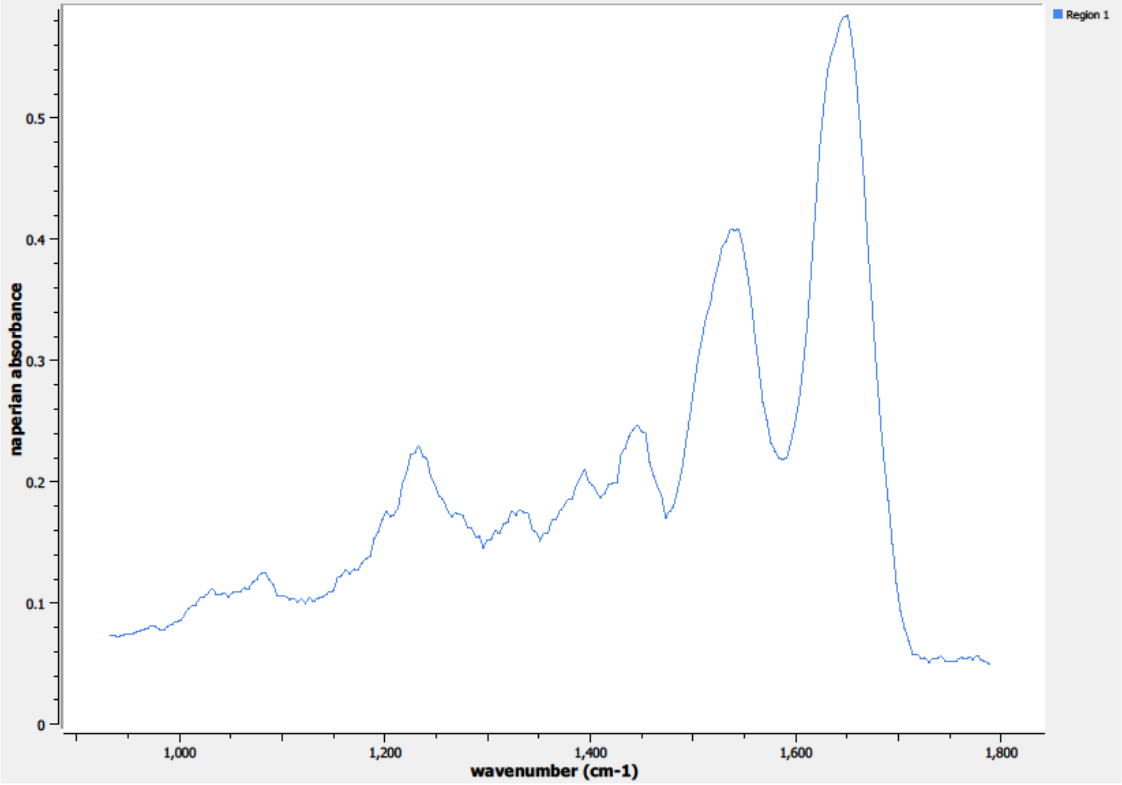



**Figure 0.42. Point I Spectra**

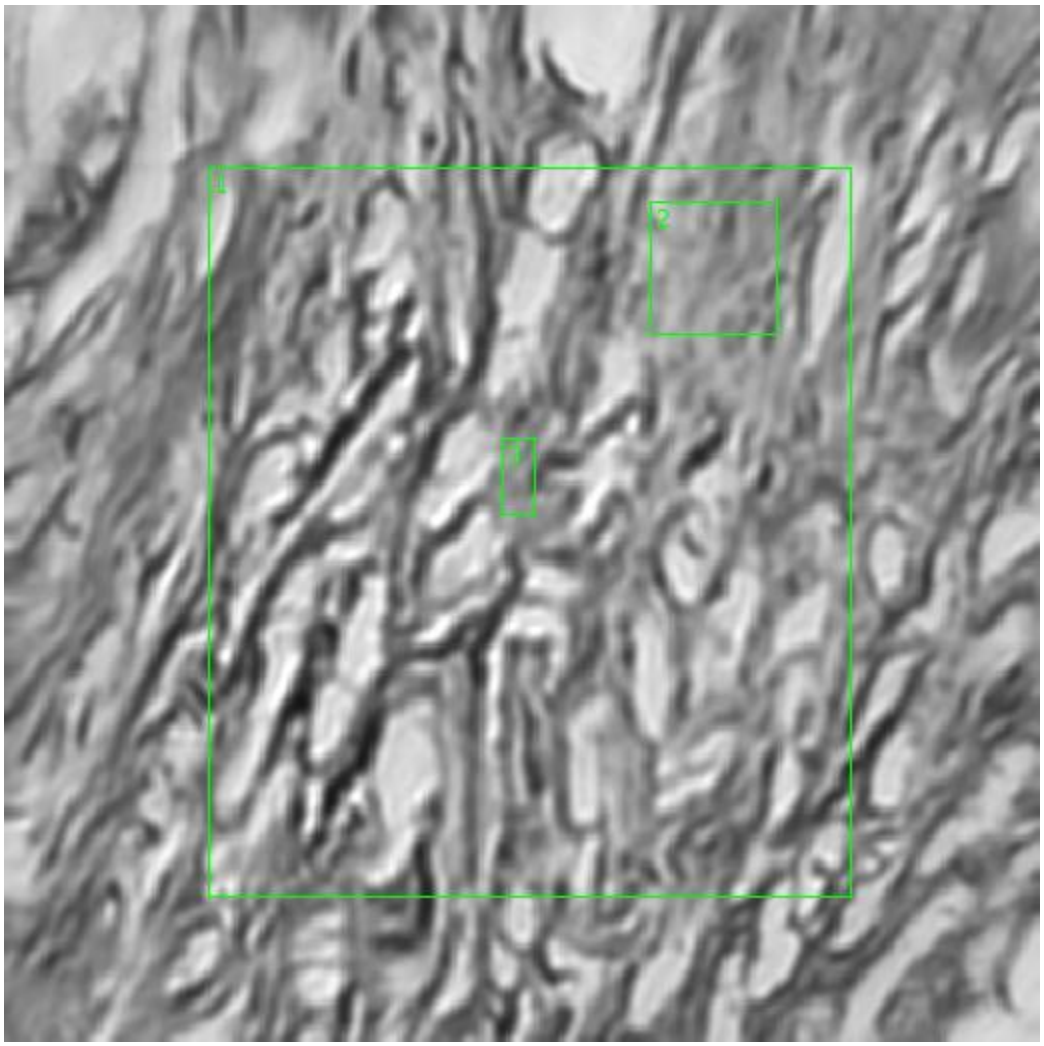



**Figure 0.43. Point J Regions of Interest**
**Region 1: breast**
**Regions 2, 3: concentrated regions, breast**

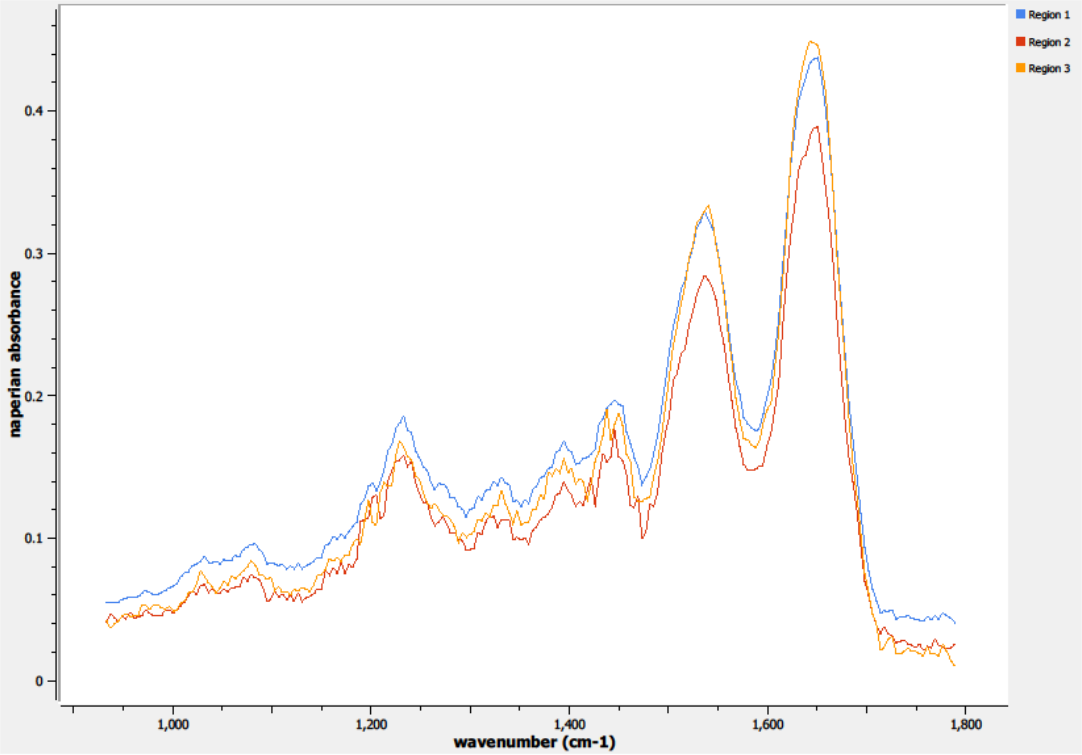



**Figure 0.44. Point J Spectra**

## A7.12-12-A057a

| | |
|---|---|
| Unique ID | 12-12-A057a |
| Age | 60 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | abnormal tissue sample, may not use |

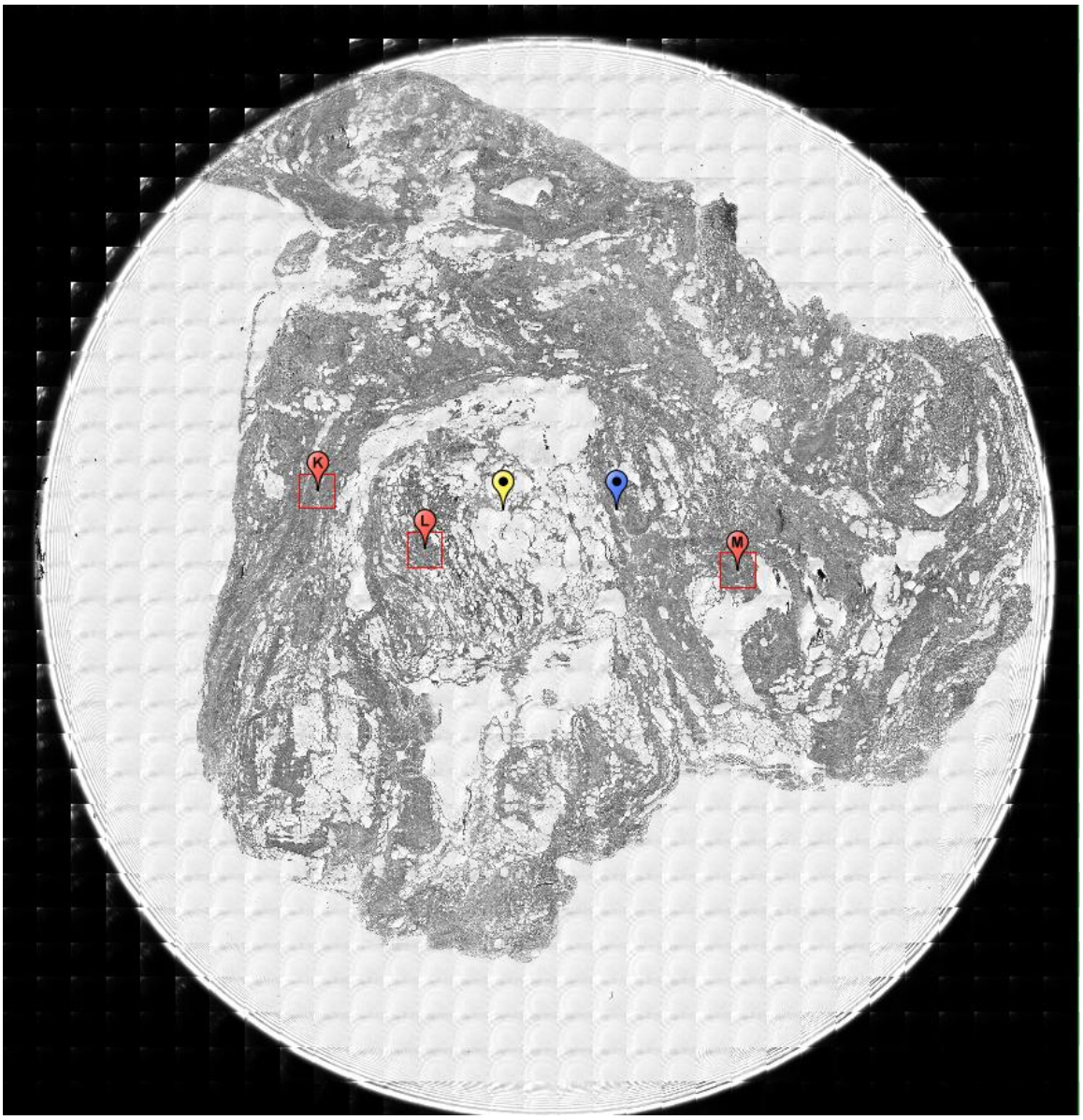



**Figure 0.45. Specimen 12-12-A057a, normal breast tissue.**

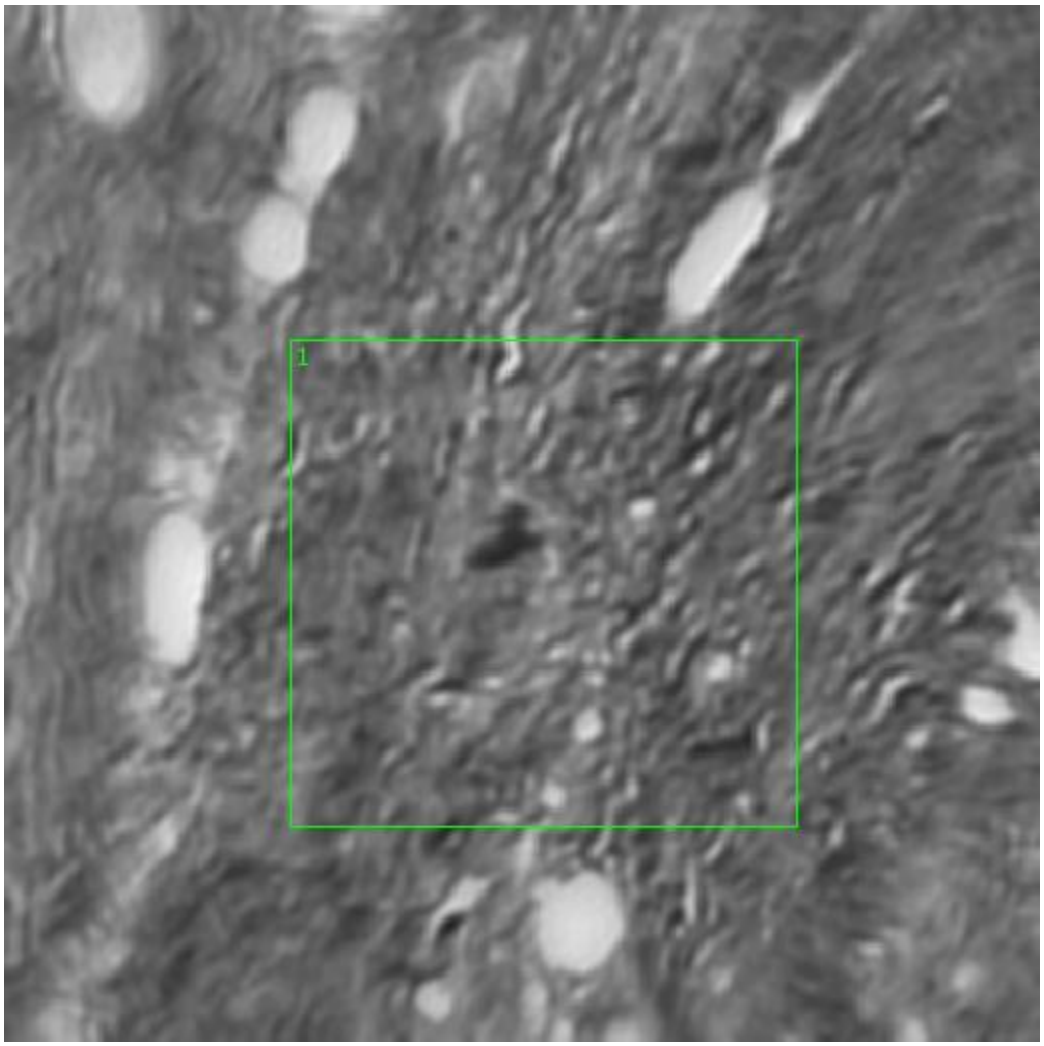



**Figure 0.46. Point K Regions of Interest**
**Region 1: breast**

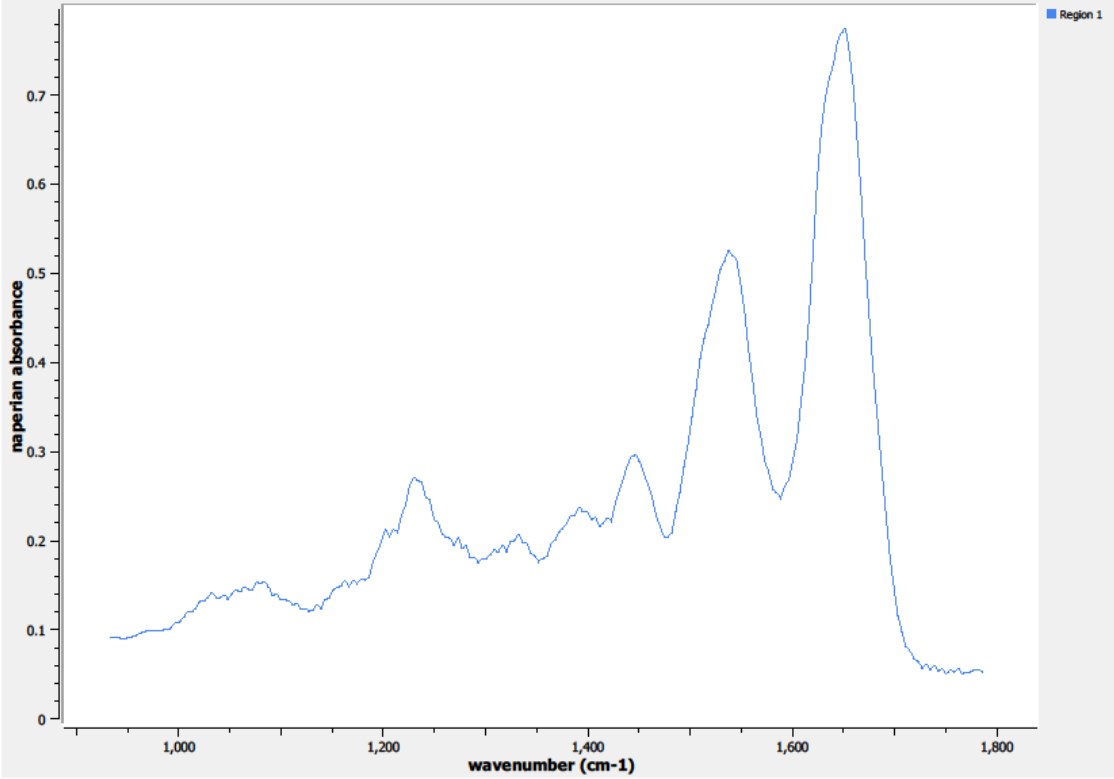



**Figure 0.47. Point K Spectra**

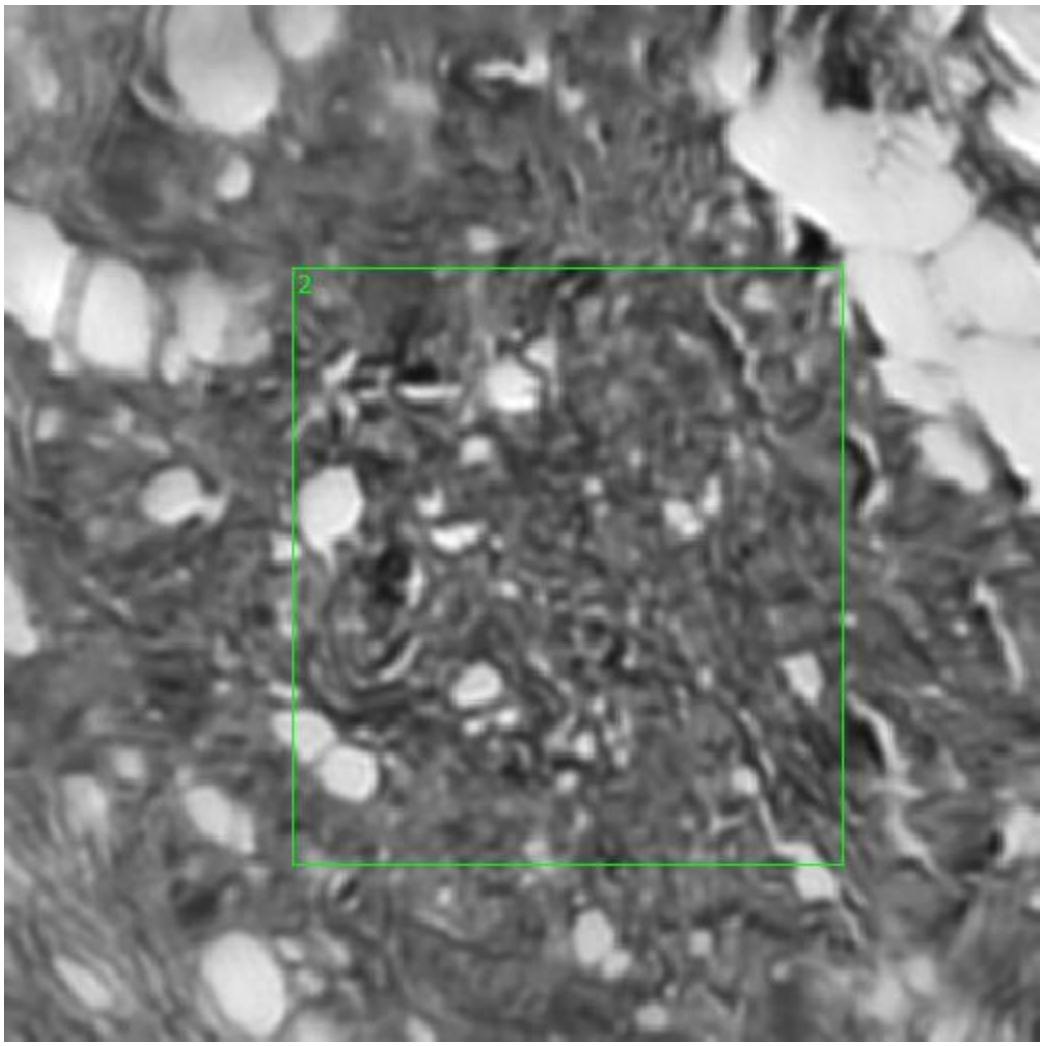



**Figure 0.48. Point L Regions of Interest**
**Region 1: breast**

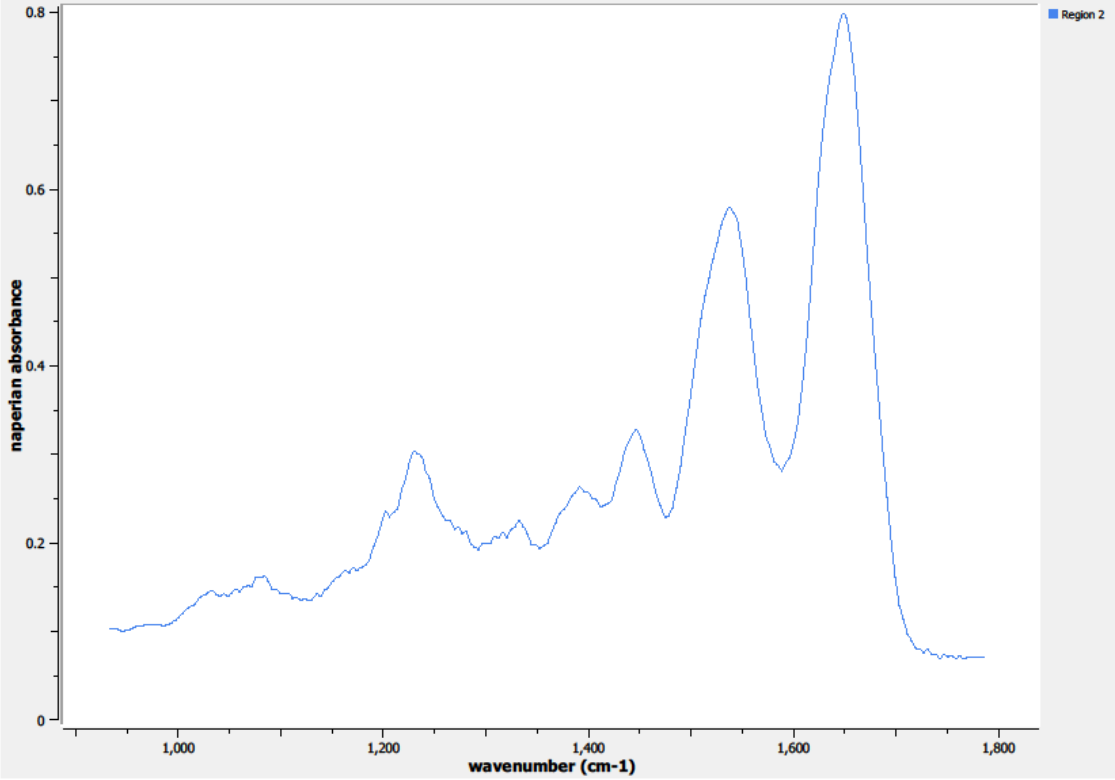



**Figure 0.49. Point L Spectra**

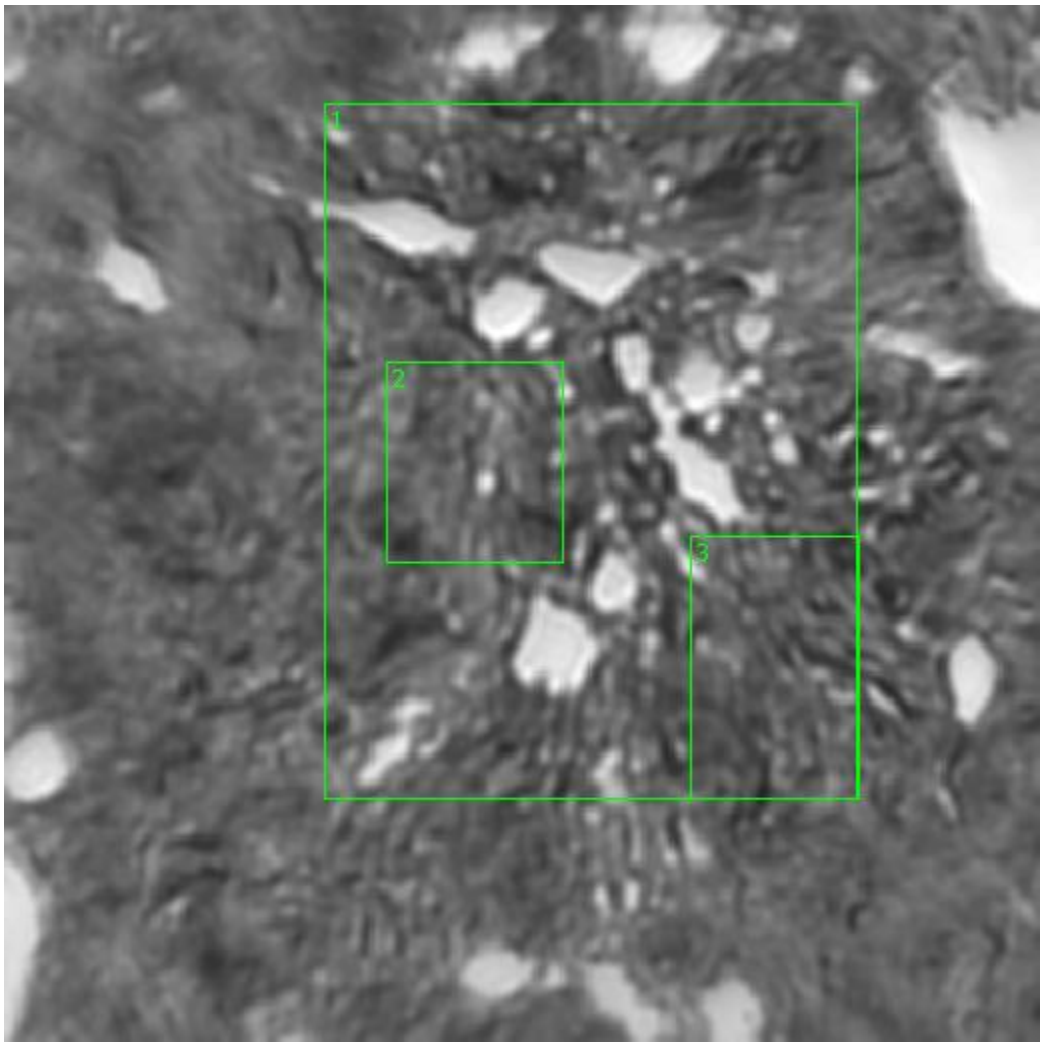



**Figure 0.50. Point M Regions of Interest**
**Region 1: breast**
**Regions 2, 3: concentrated regions, breast**

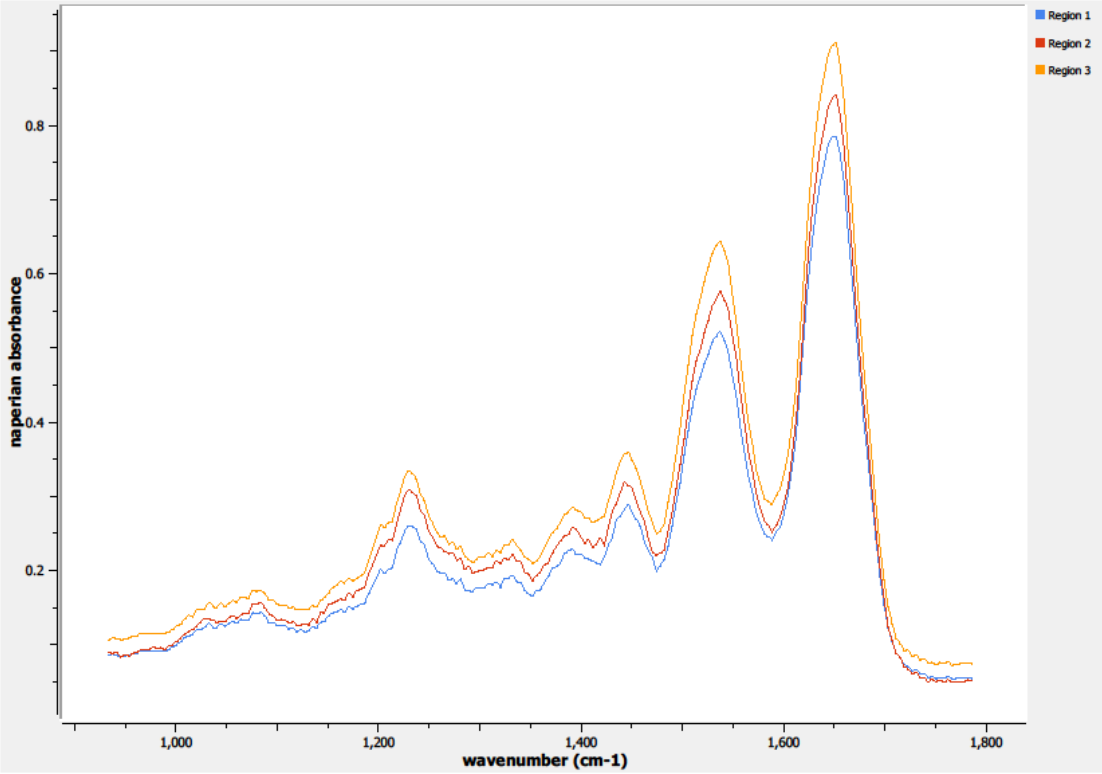



**Figure 0.51. Point M Spectra**

## A8.12-12-A095a

| | |
|---|---|
| Unique ID | 12-12-A095a |
| Age | 28 |
| Race | white |
| Location of Sample Collection | right breast |
| Notes of Interest | none |

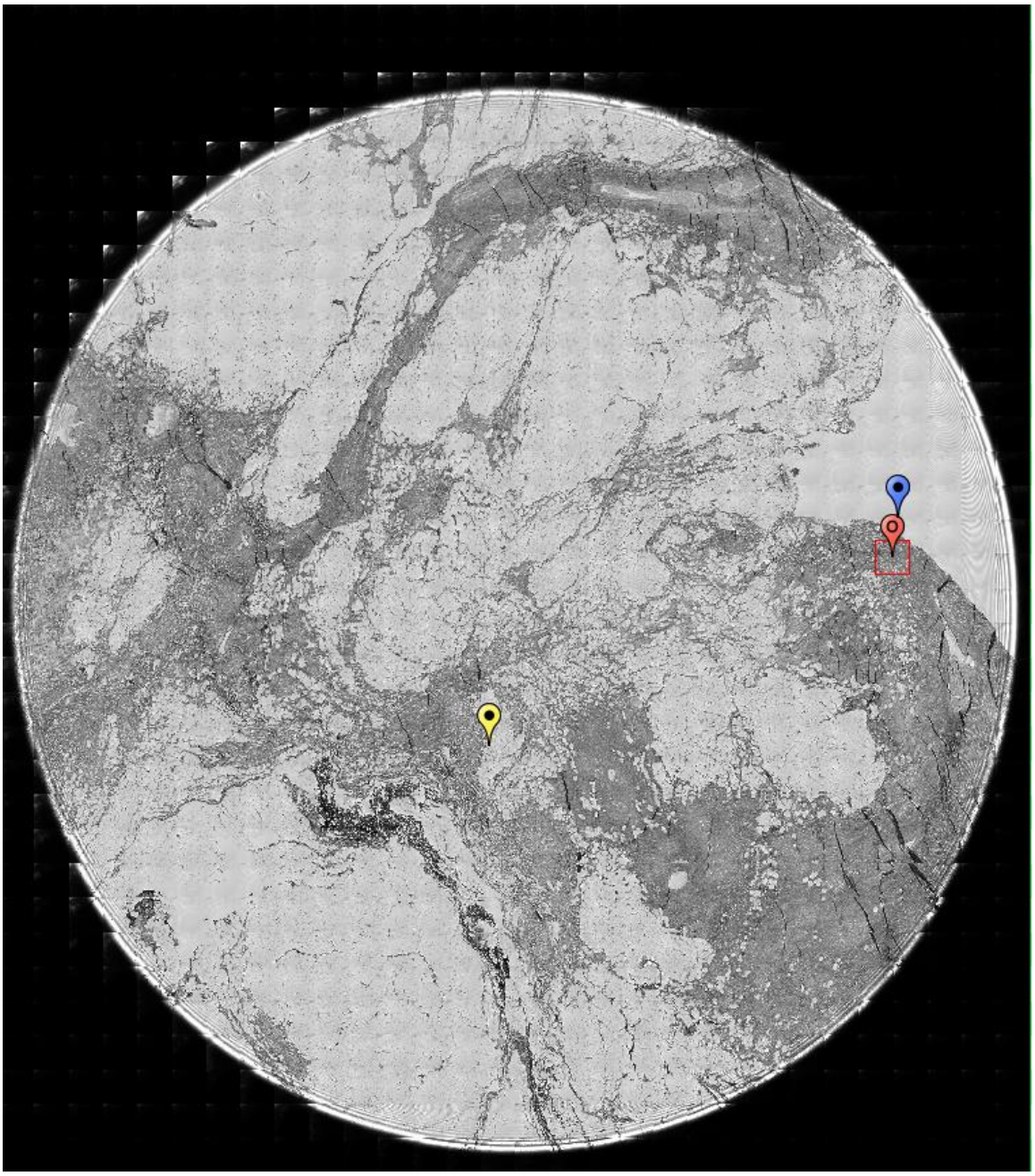


**Figure 0.52. Specimen 12-12-A095a, normal right breast tissue.**

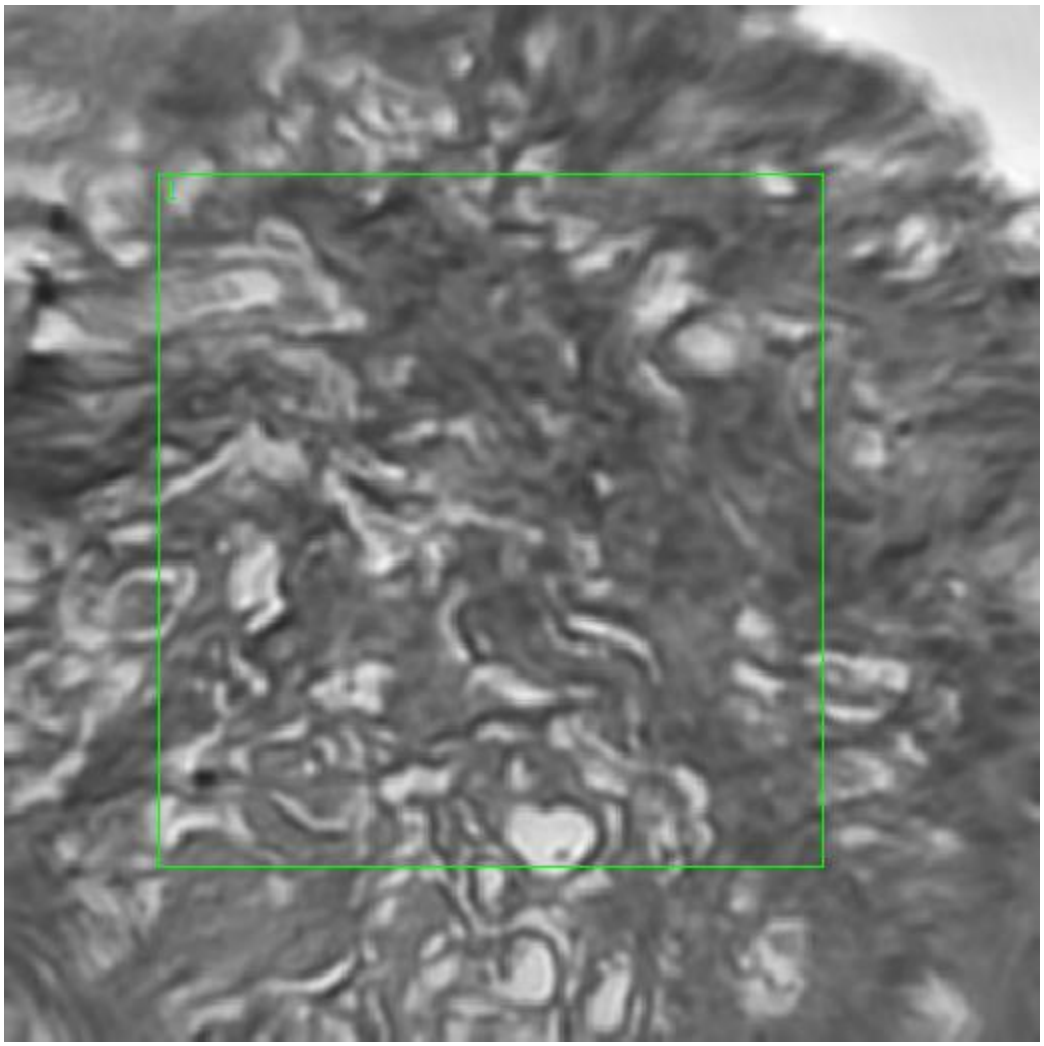


**Figure 0.53. Point O Regions of Interest**
**Region 1:  normal right breast tissue**

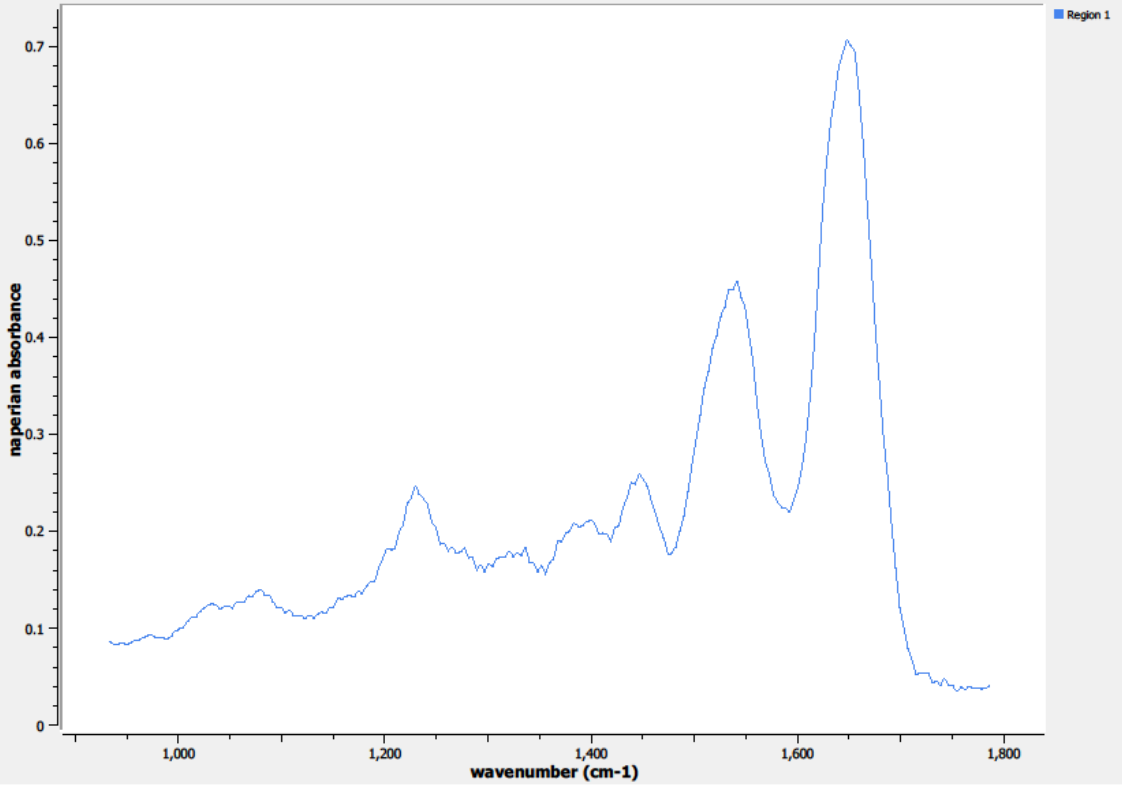



**Figure 0.54. Point O Spectra**

## A9.12-12-A099a

| | |
|---|---|
| Unique ID | 12-12-A099a |
| Age | 60 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

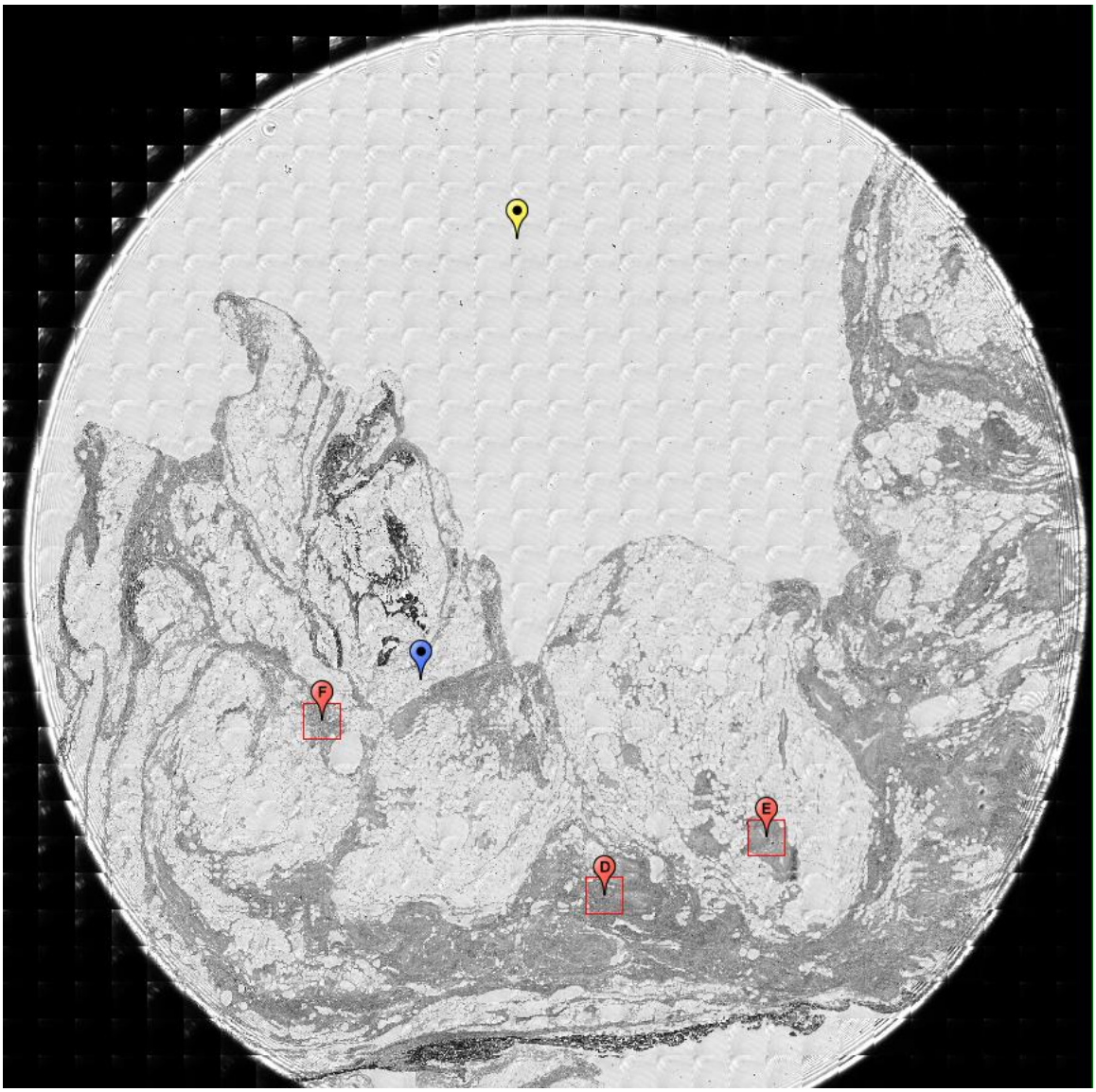



**Figure 0.55. Specimen 12-12-A099a, normal breast tissue.**

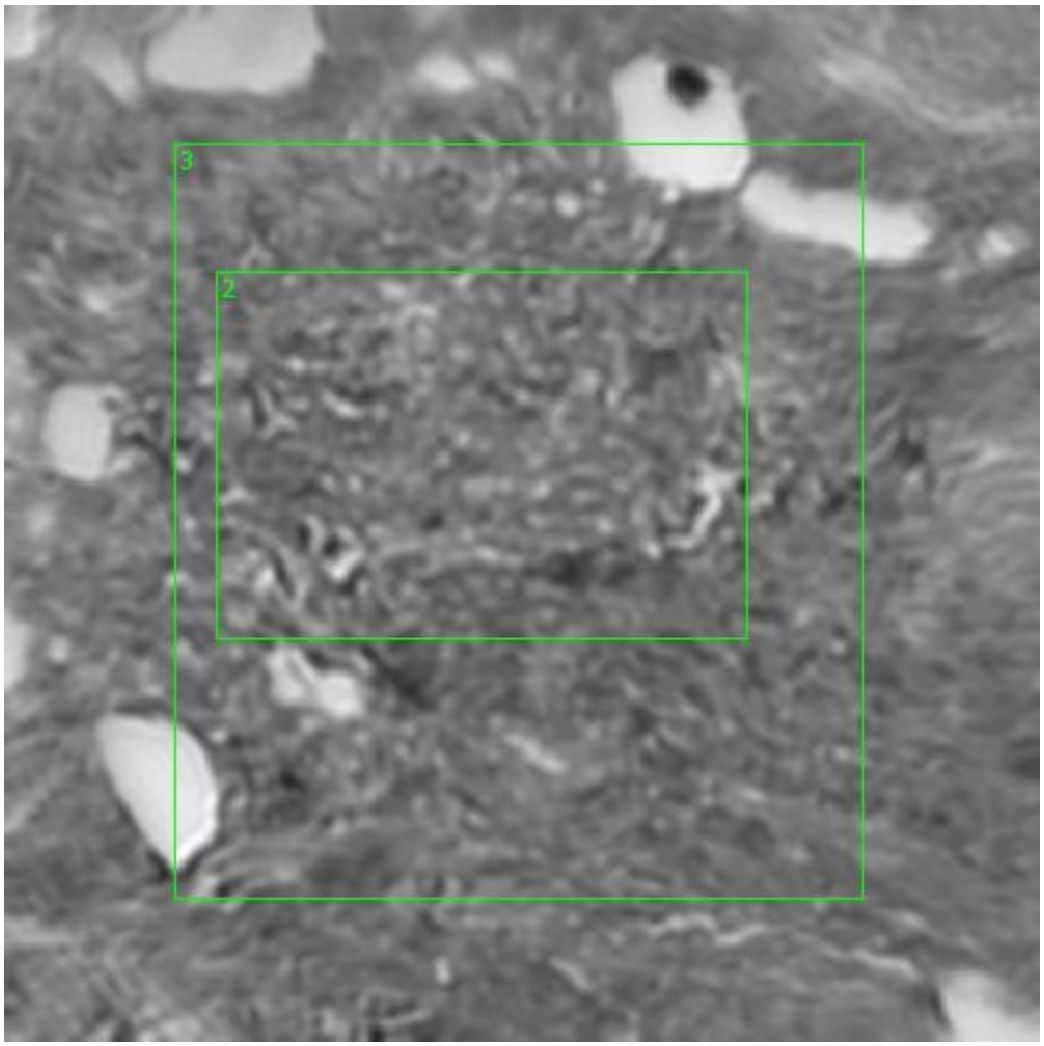



**Figure 0.56. Point D Regions of Interest**
**Regions 2, 3: breast**

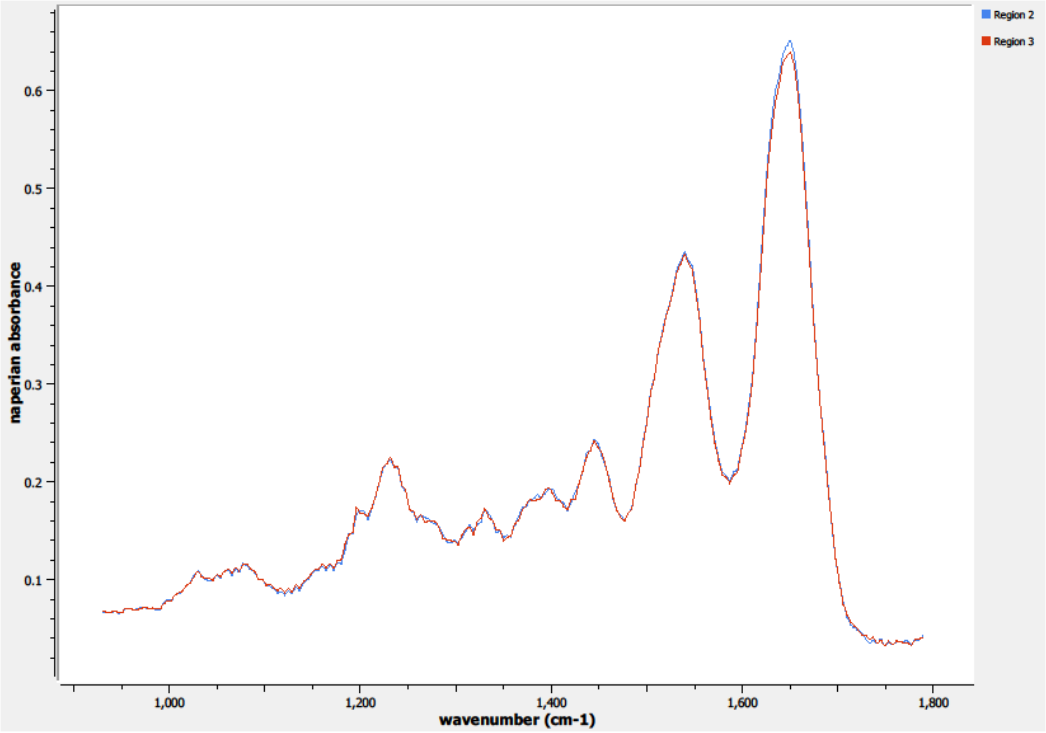



**Figure 0.57. Point D Spectra**

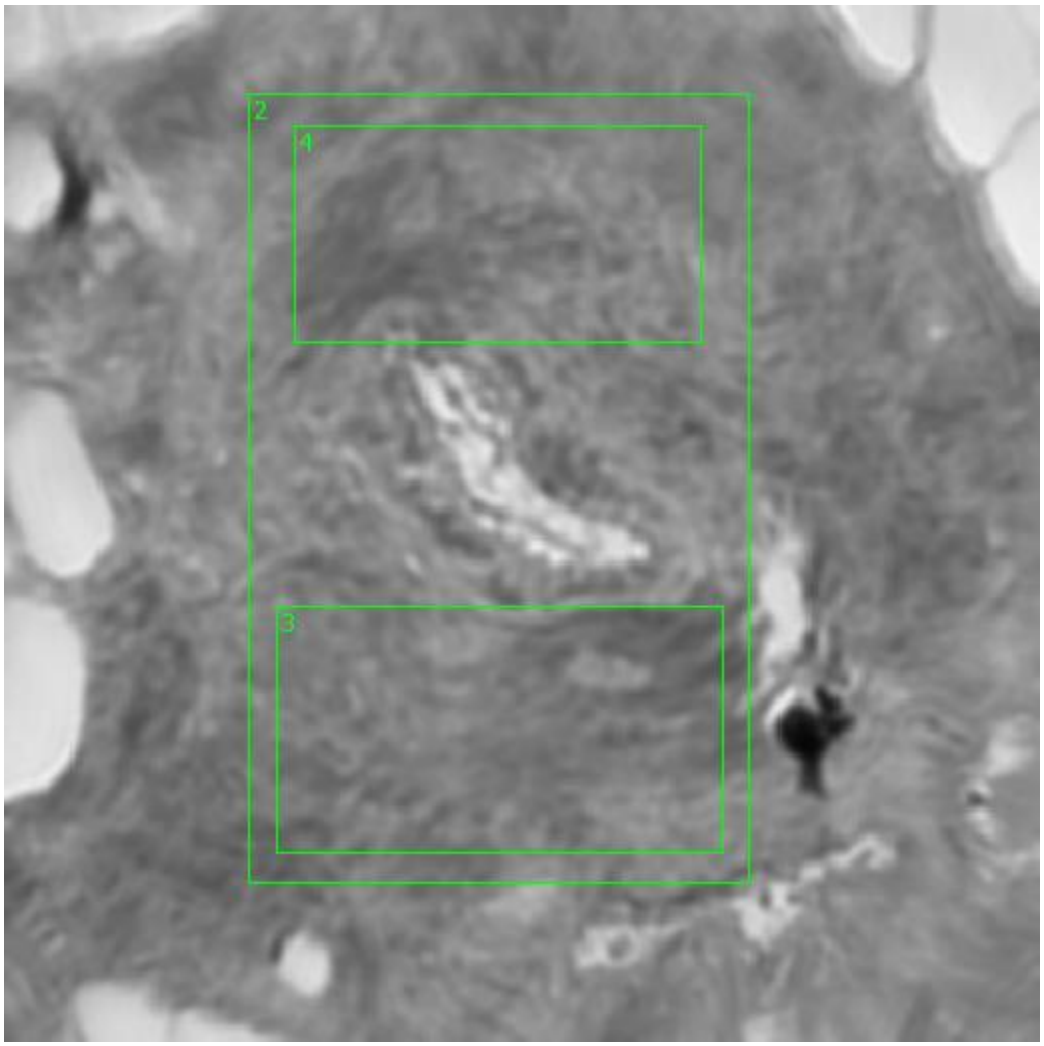



**Figure 0.58. Point E Regions of Interest**
**Region 2: breast**
**Regions 3, 4: concentrated regions, breast**

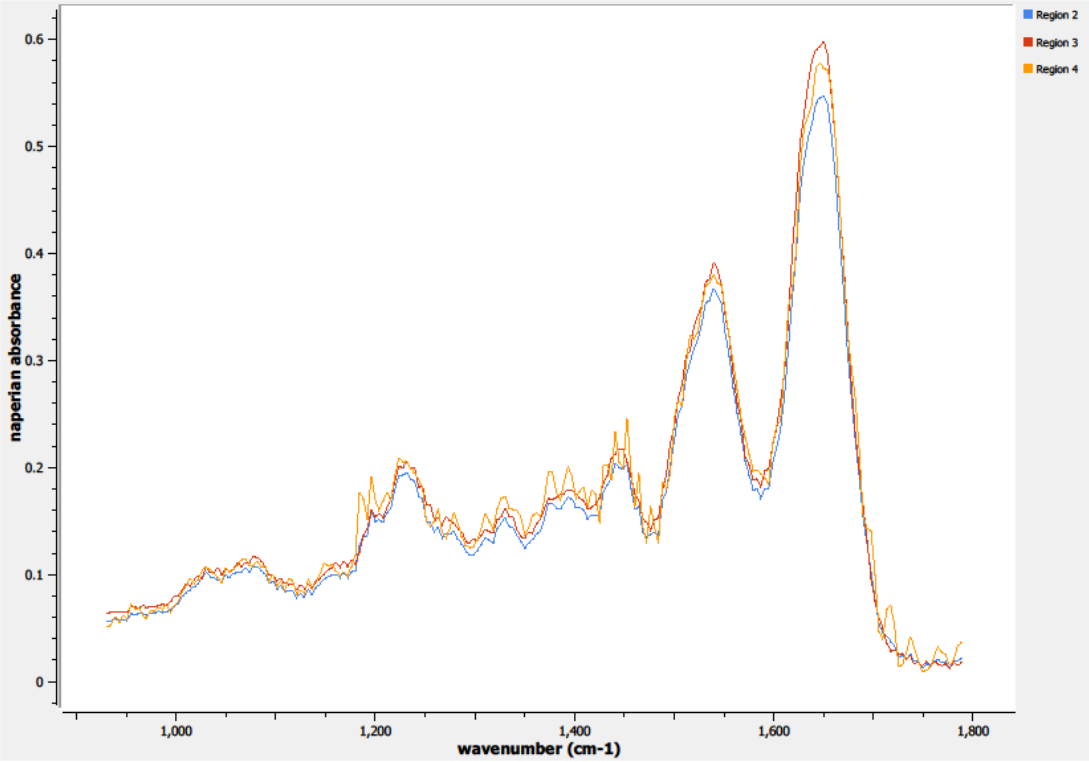



**Figure 0.59. Point E Spectra**

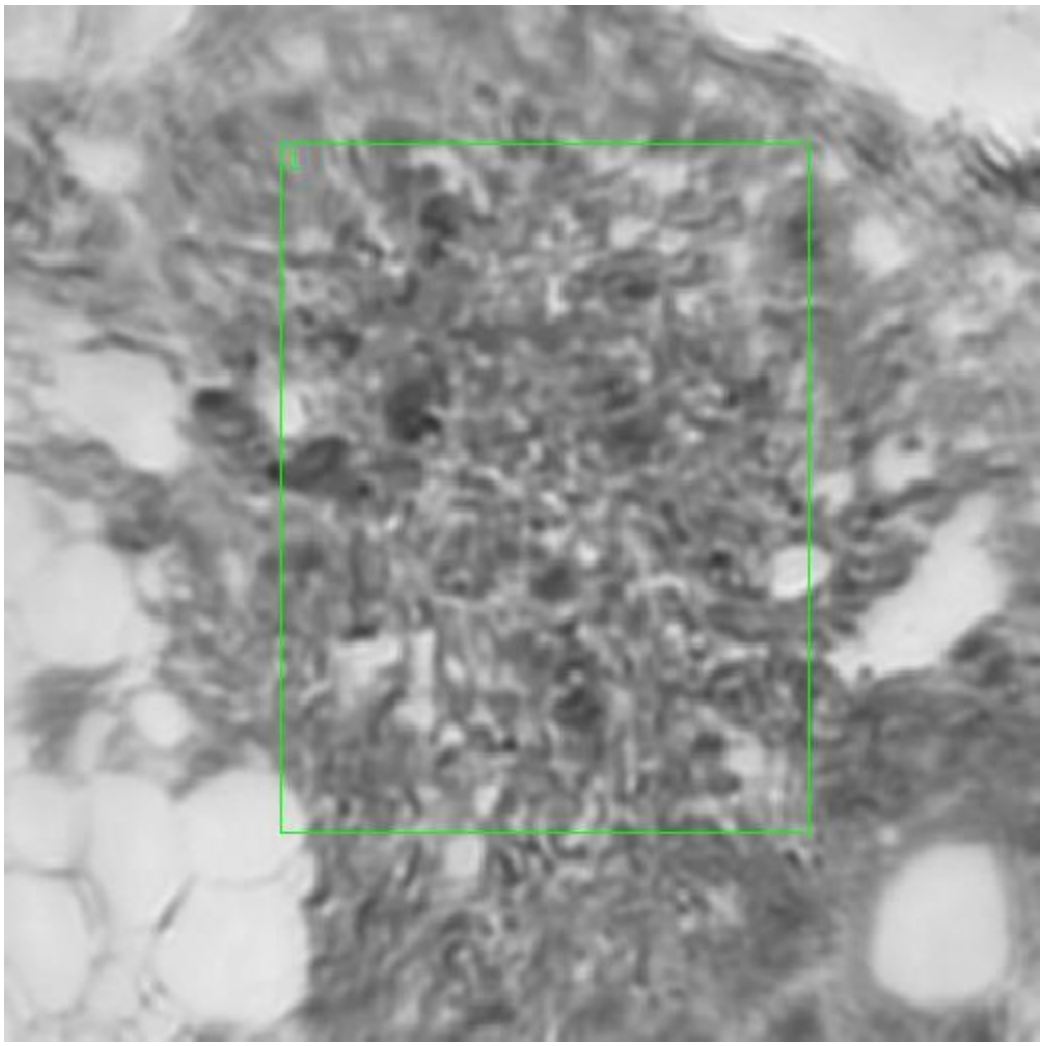


**Figure 0.60. Point F Regions of Interest**
**Region 1: breast**

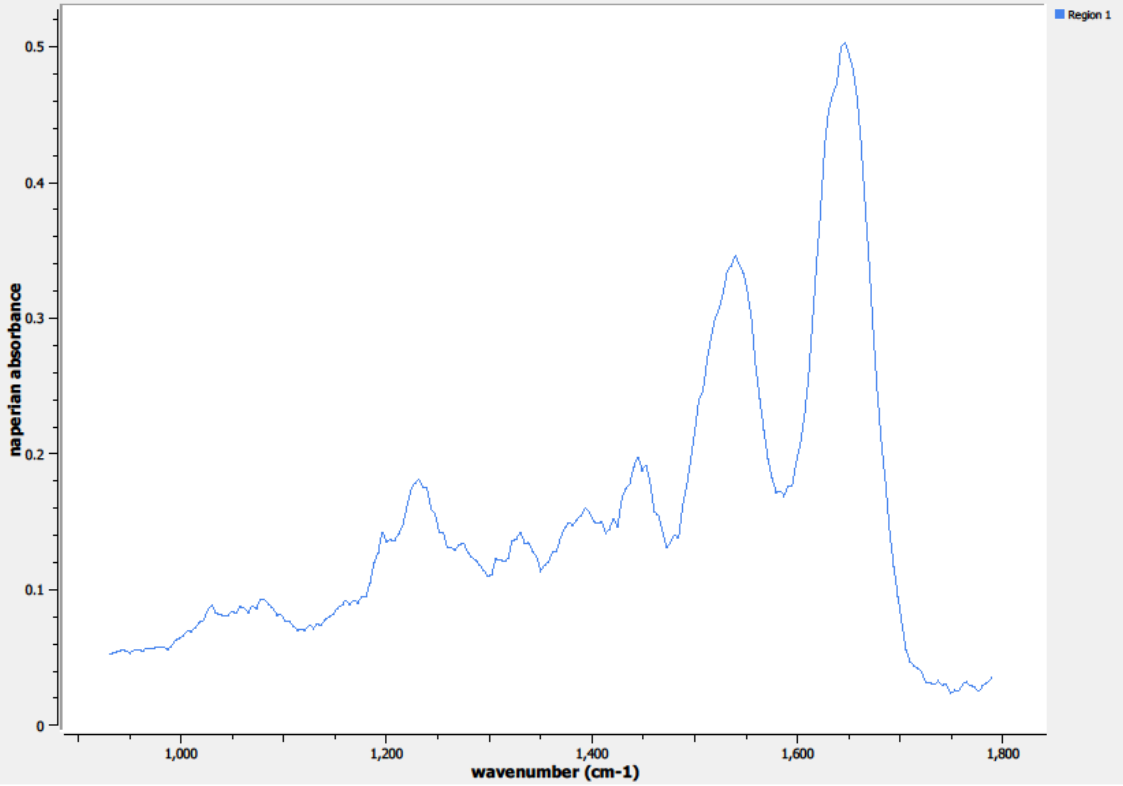



**Figure 0.61. Point F Spectra**

## A10. 12-12-A122a

| | |
|---|---|
| Unique ID | 12-12-A122a |
| Age | 58 |
| Race | white |
| Location of Sample Collection | right breast |
| Notes of Interest | none |

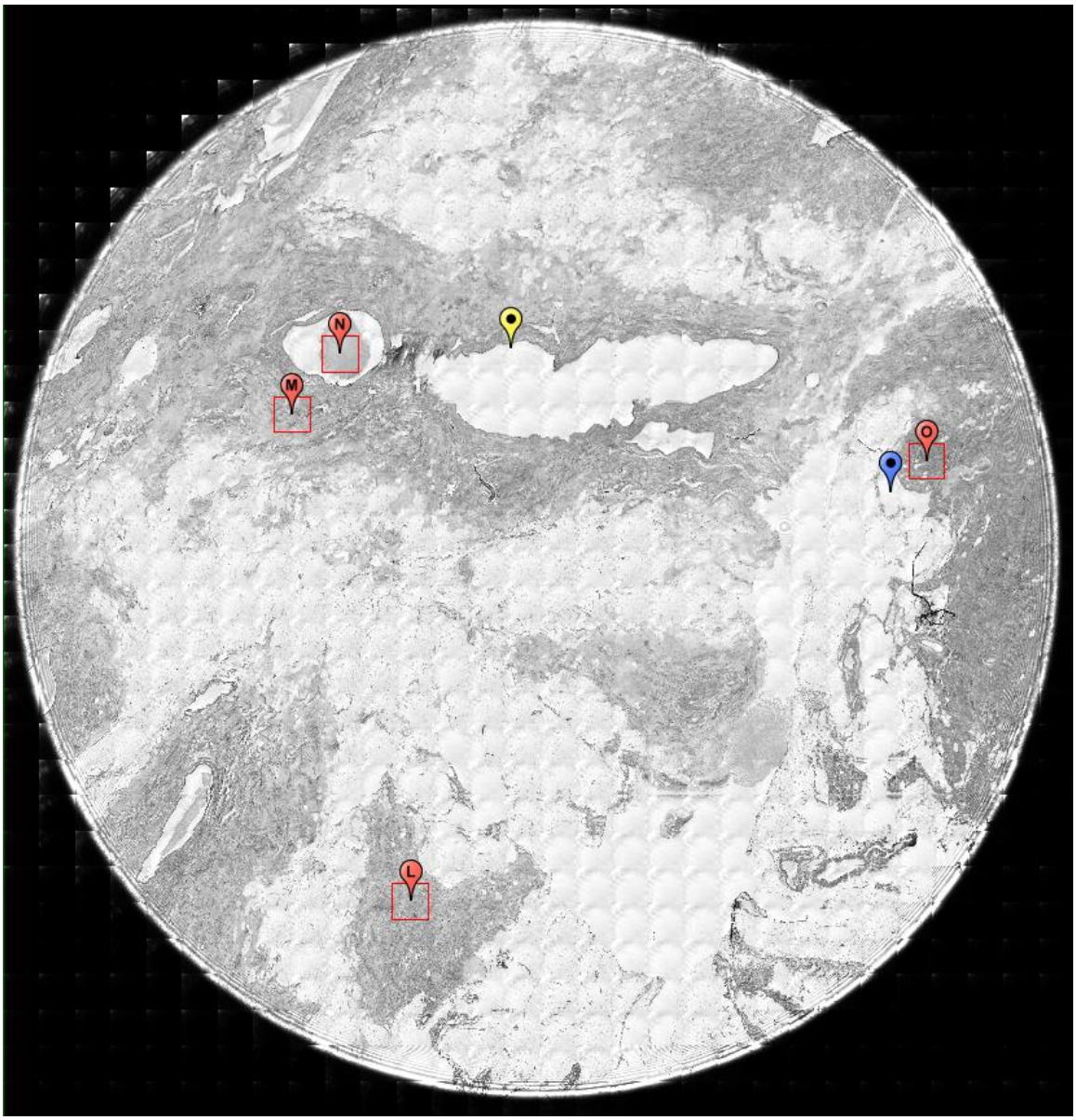



**Figure 0.62. Specimen 12-12-A122a, normal right breast tissue.**

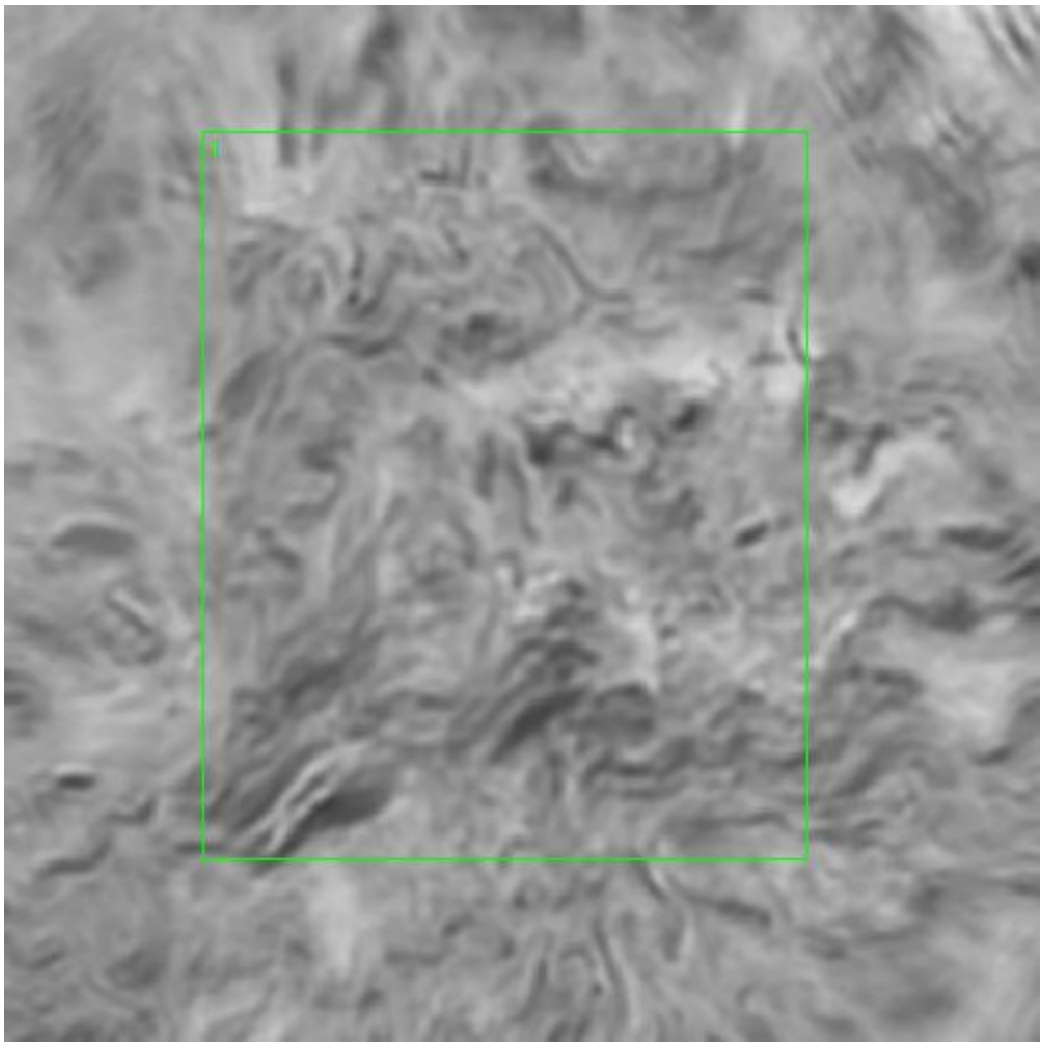


**Figure 0.63. Point L Regions of Interest**
**Region 1: right breast**

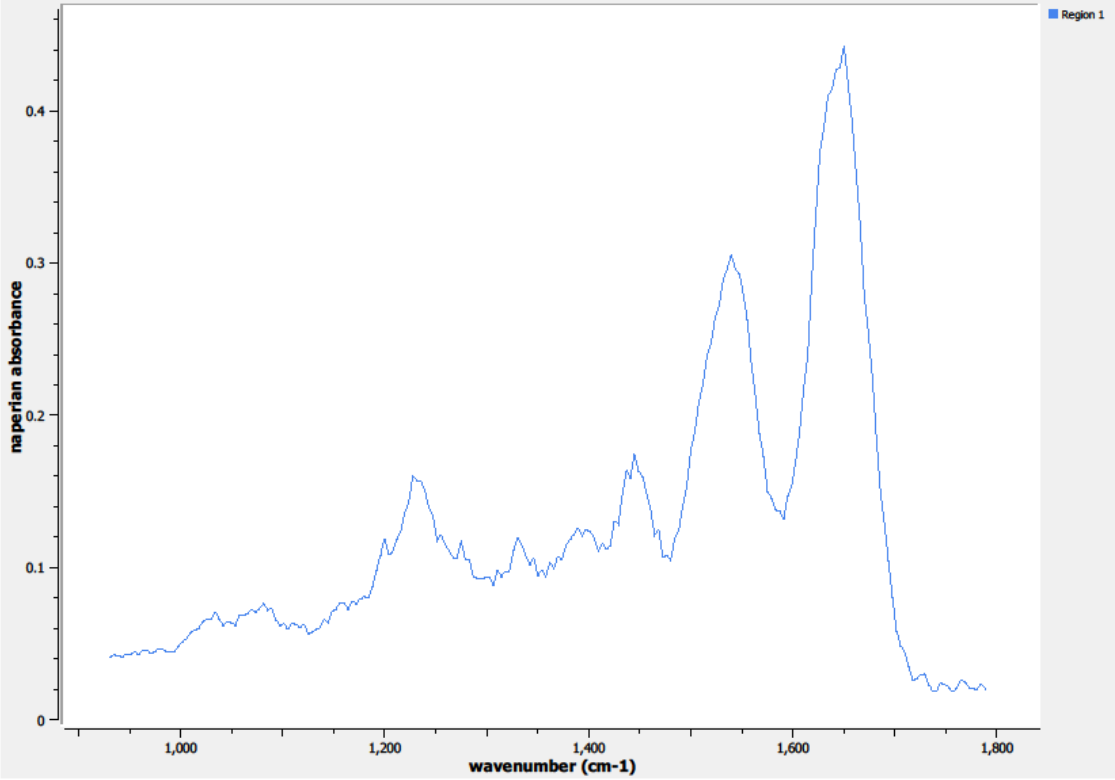



**Figure 0.64. Point L Spectra**

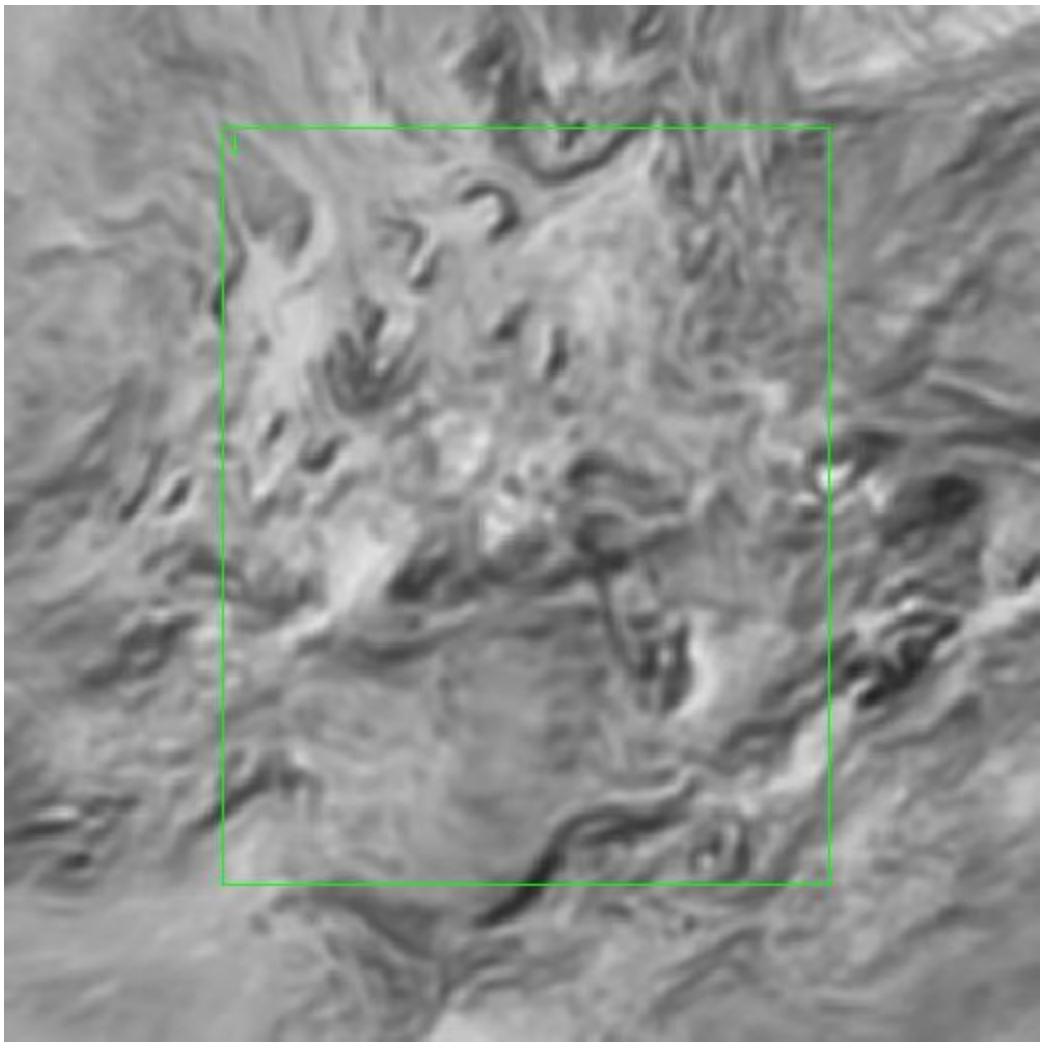


**Figure 0.65. Point M Regions of Interest**
**Region 1: right breast**

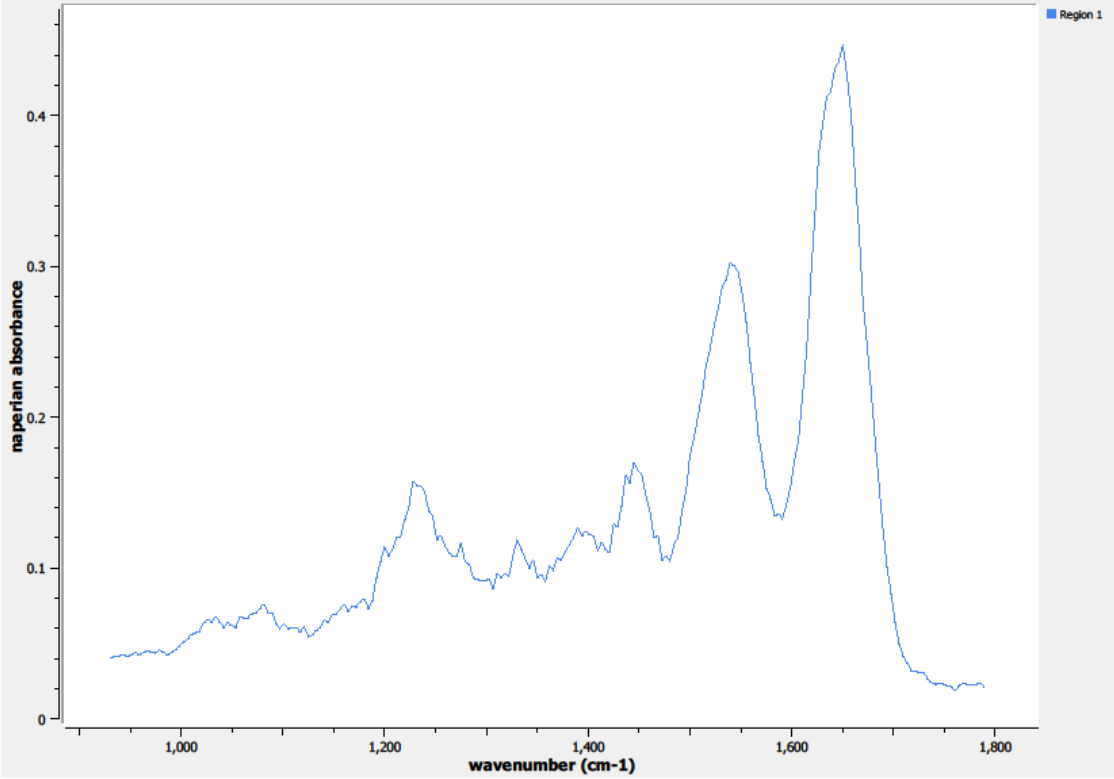



**Figure 0.66. Point M Spectra**

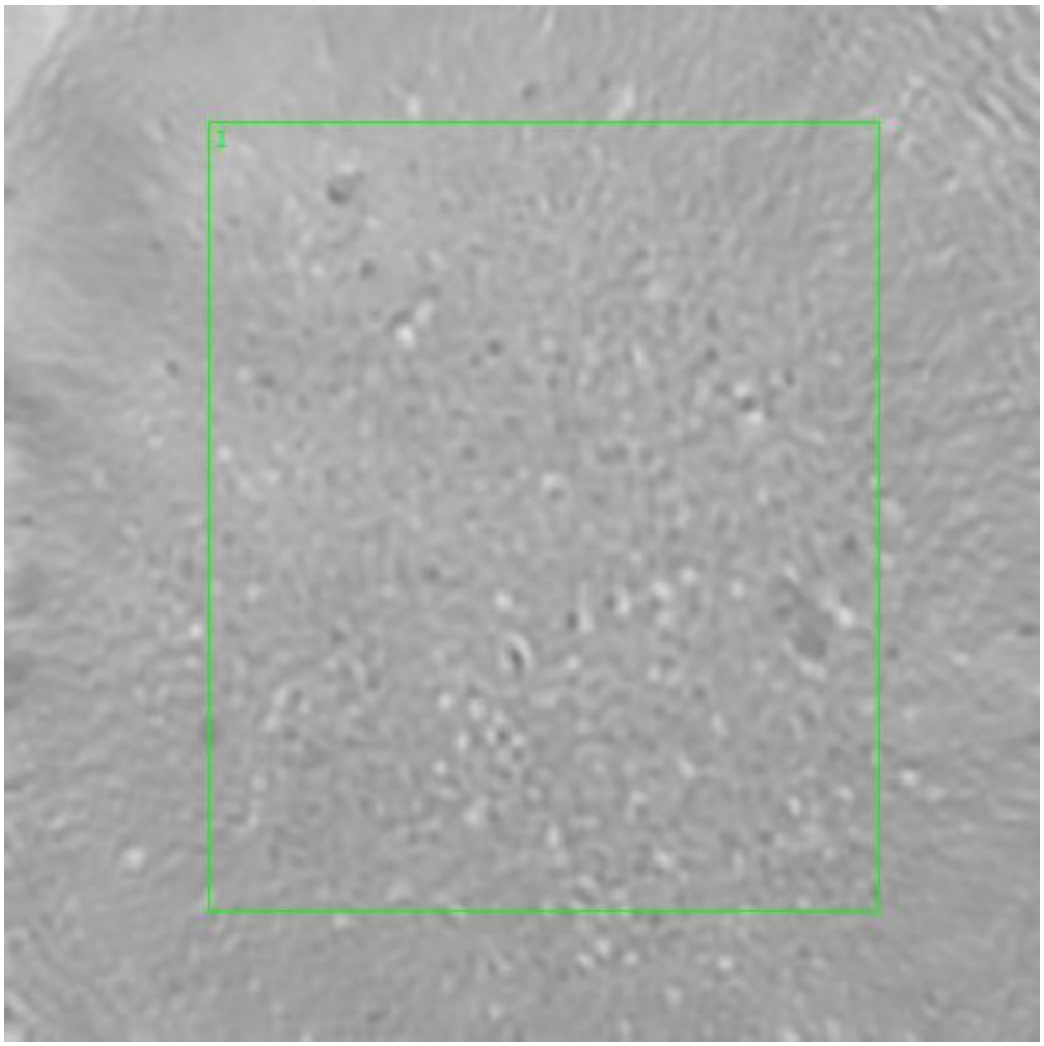


**Figure 0.67. Point N Regions of Interest**
**Region 1: interior breast tissue, duct**

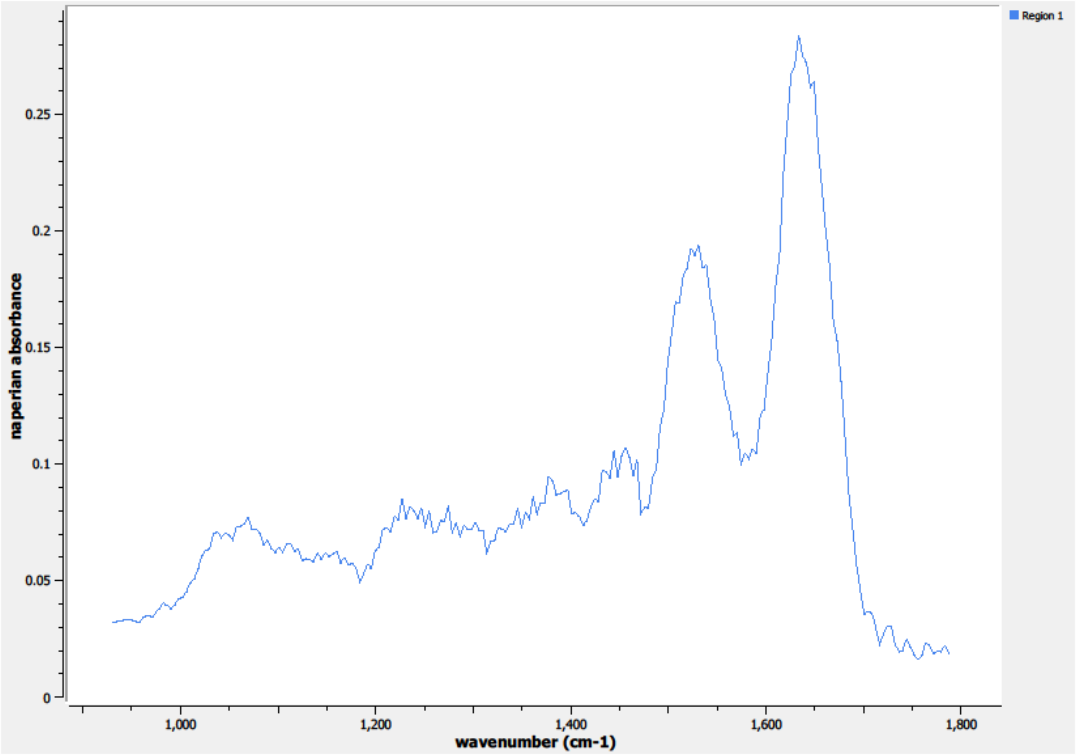



**Figure 0.68. Point N Spectra**

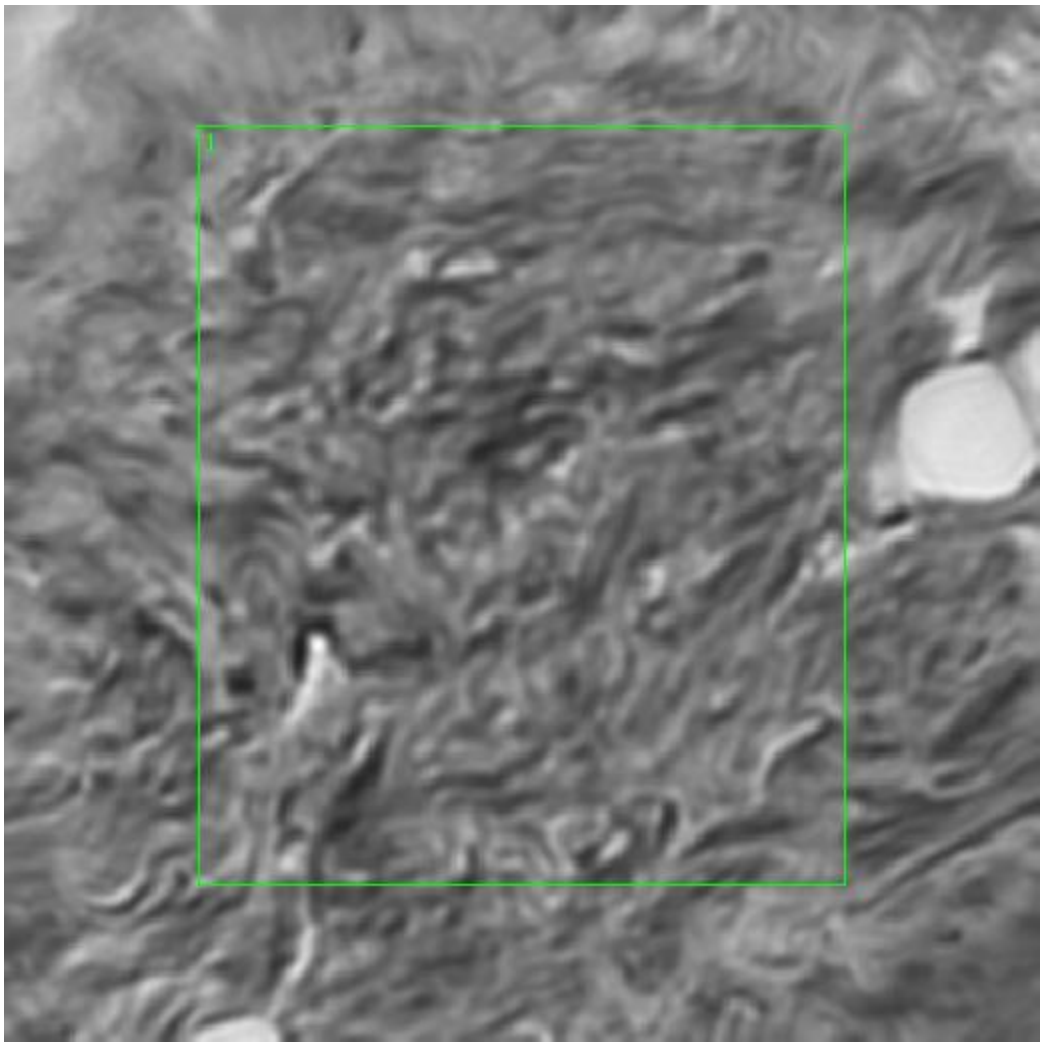


**Figure 0.69. Point O Regions of Interest**
**Region 1: right breast**

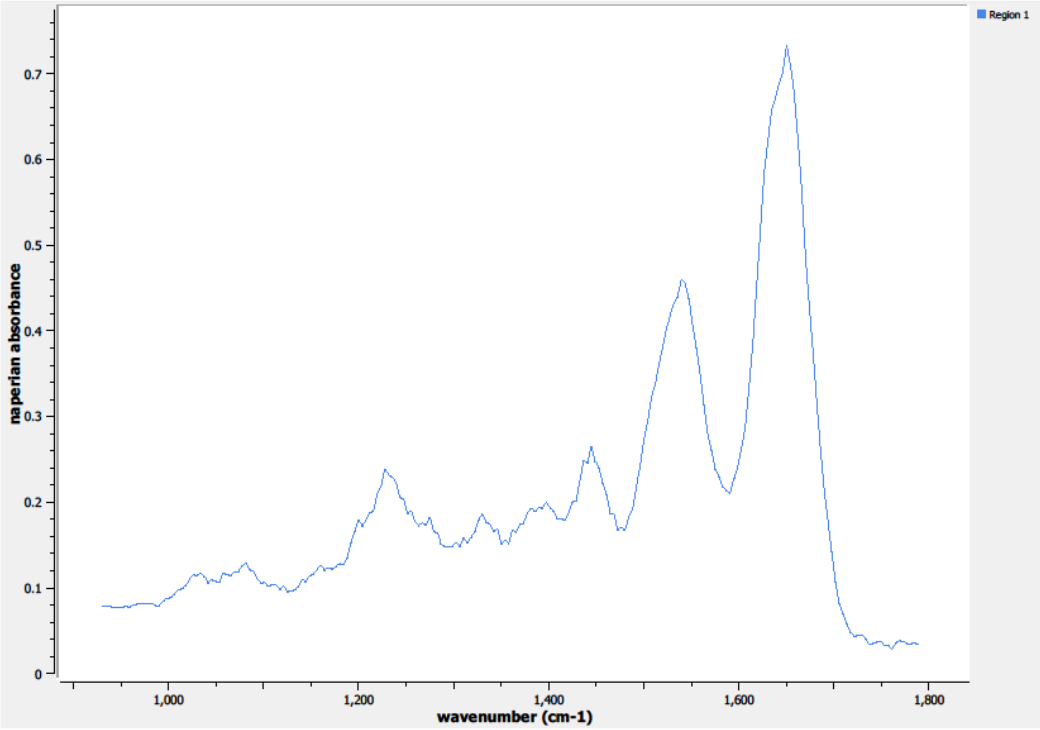



**Figure 0.70. Point O Spectra**

## A11. 13-02-A073a

| | |
|---|---|
| Unique ID | 13-02-A073a |
| Age | 59 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

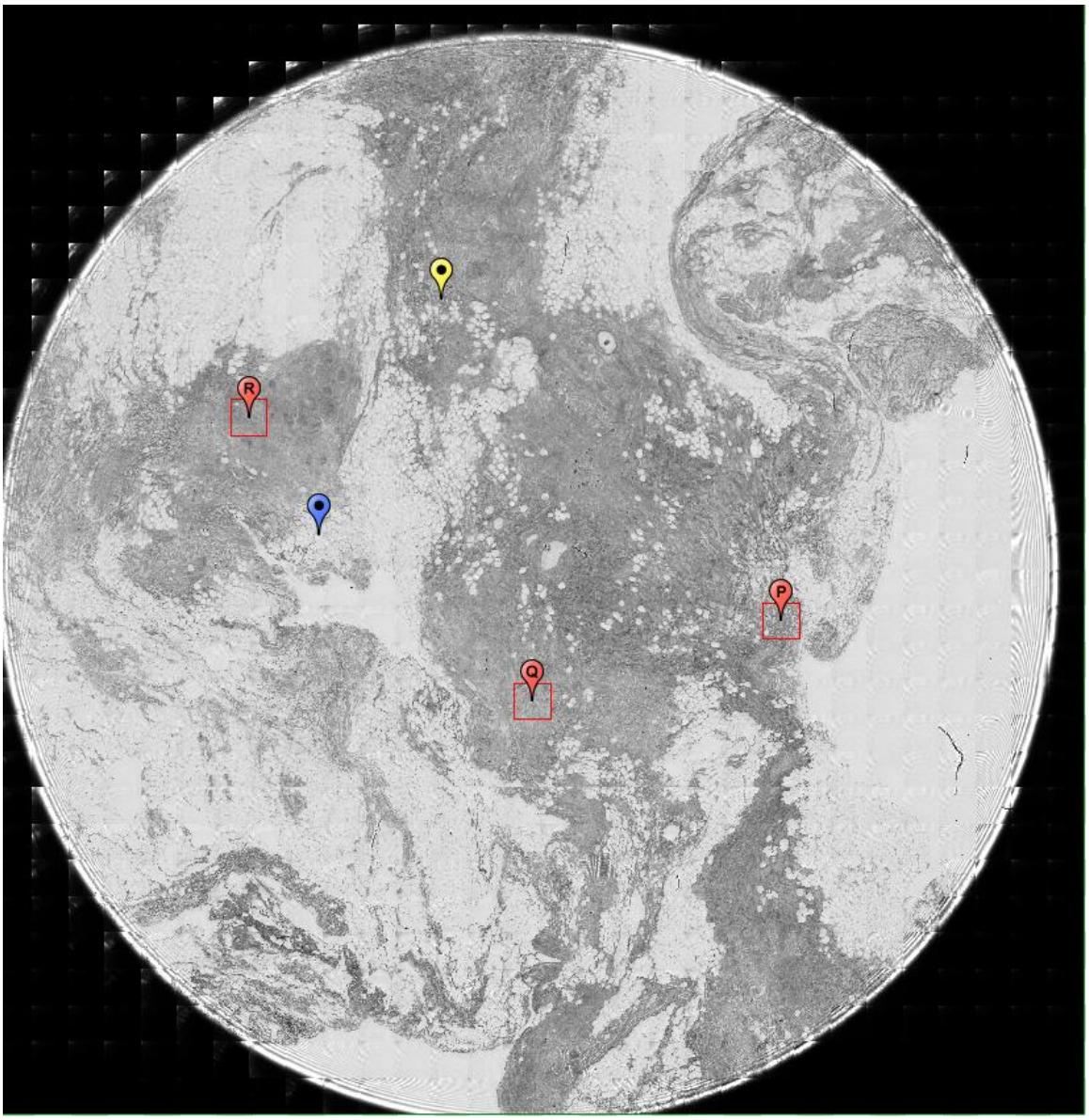



**Figure 0.71. Specimen 13-02-A073a, normal breast tissue.**

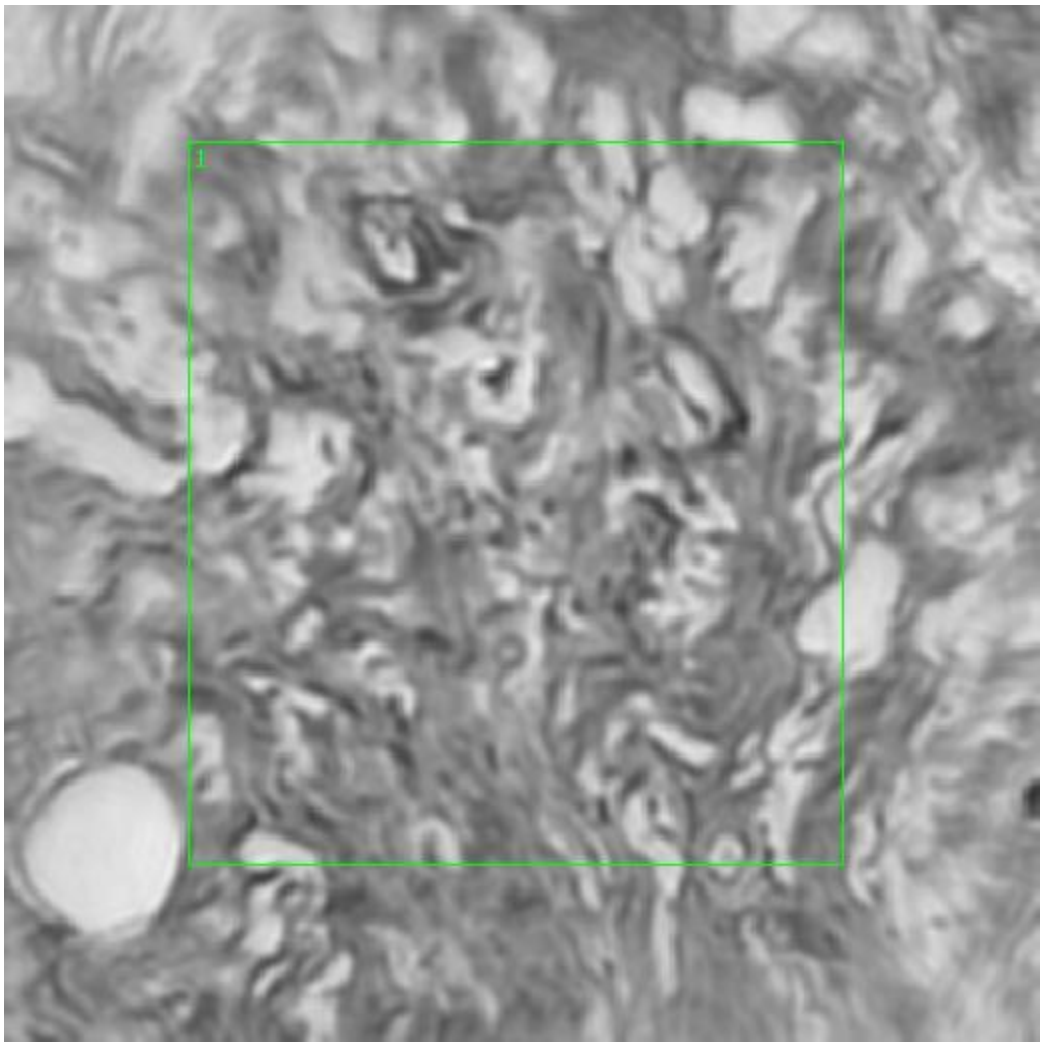



**Figure 0.72. Point P Regions of Interest**
**Region 1: breast**

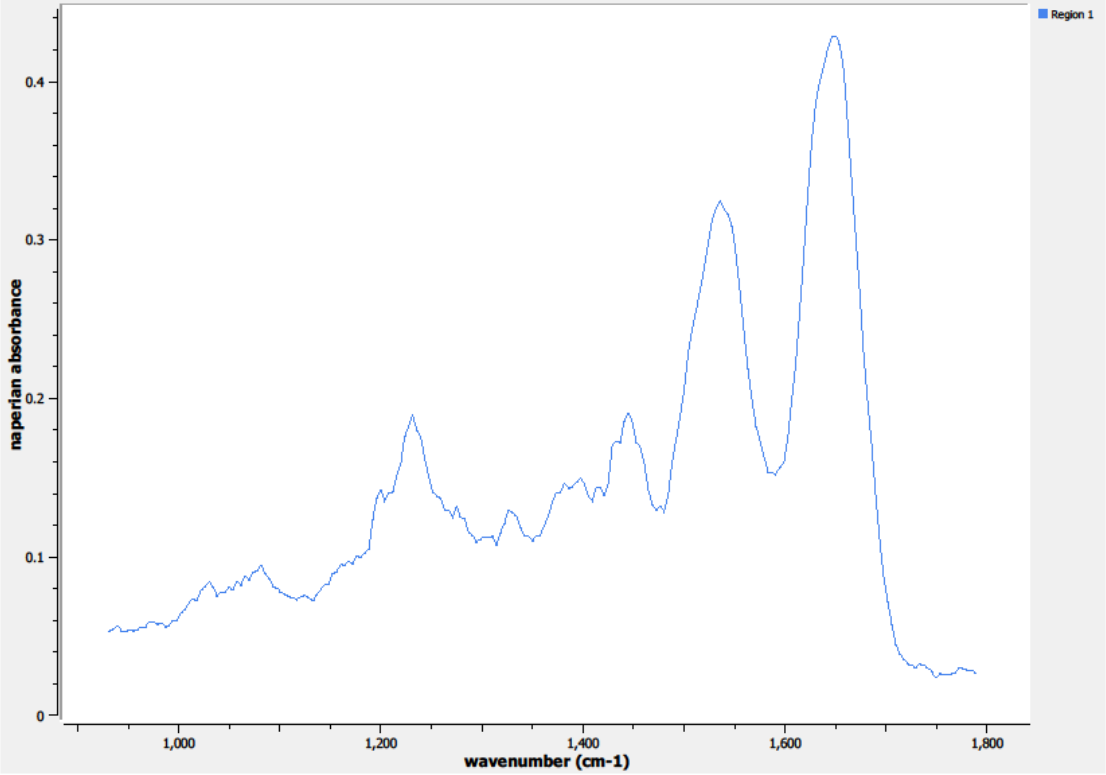



**Figure 0.73. Point P Spectra**

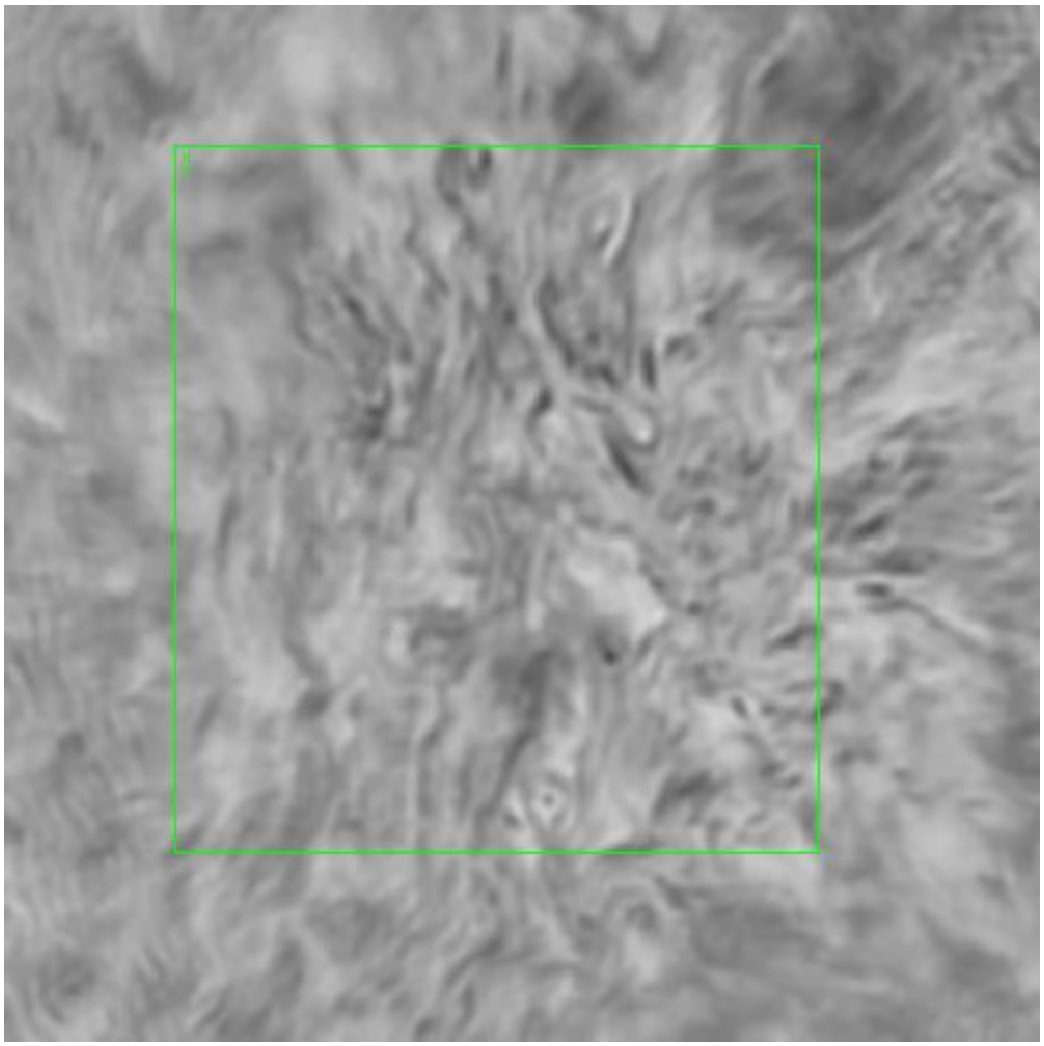


**Figure 0.74. Point Q Regions of Interest**
**Region 1: breast**

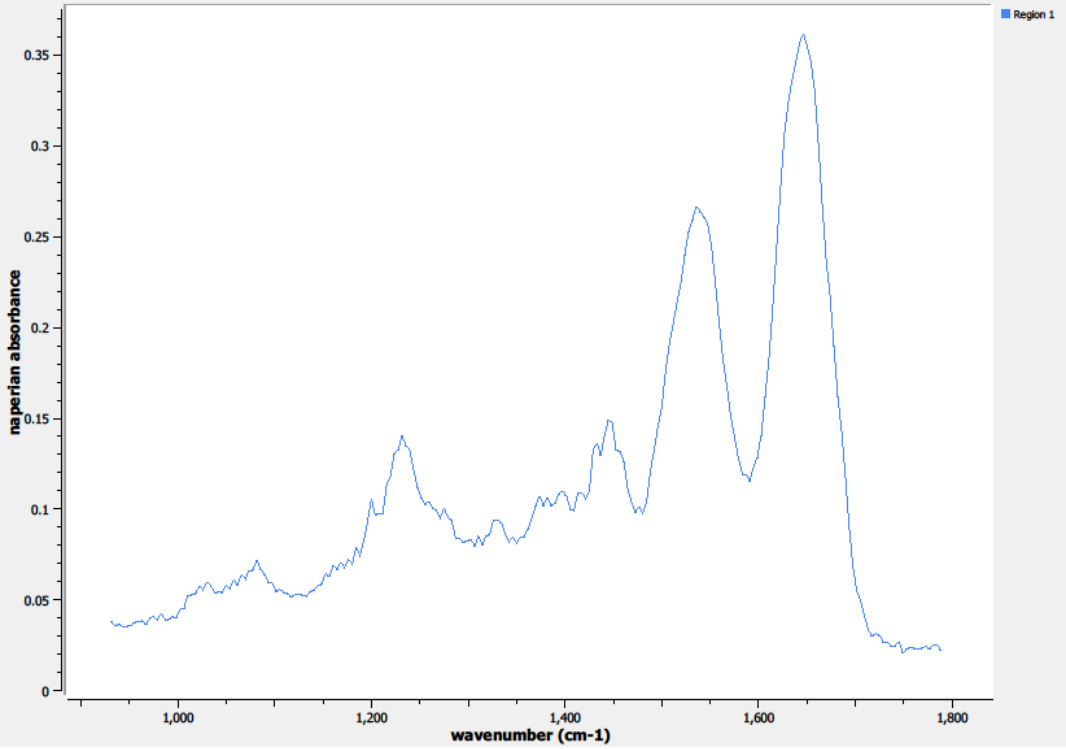



**Figure 0.75. Point Q Spectra**

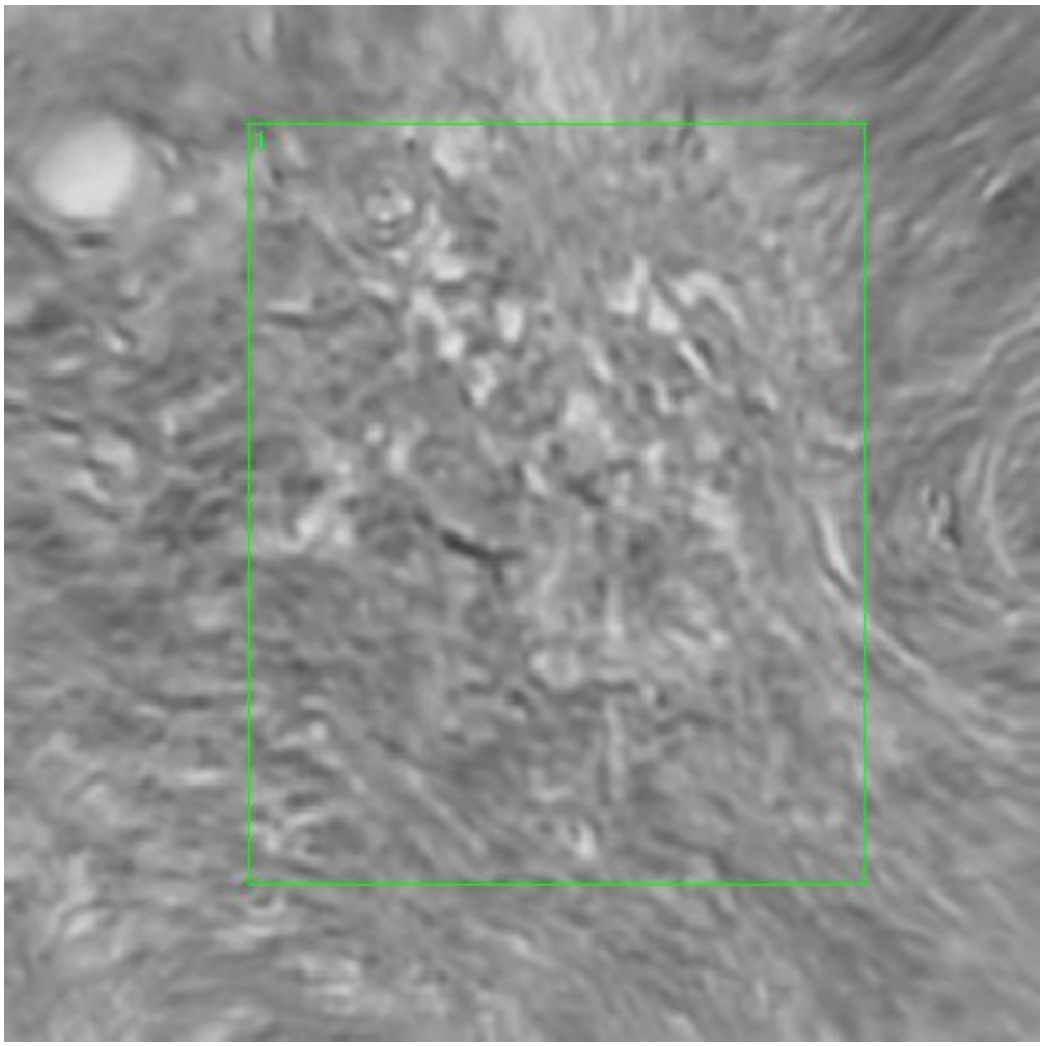


**Figure 0.76. Point R Regions of Interest**
**Region 1: breast**

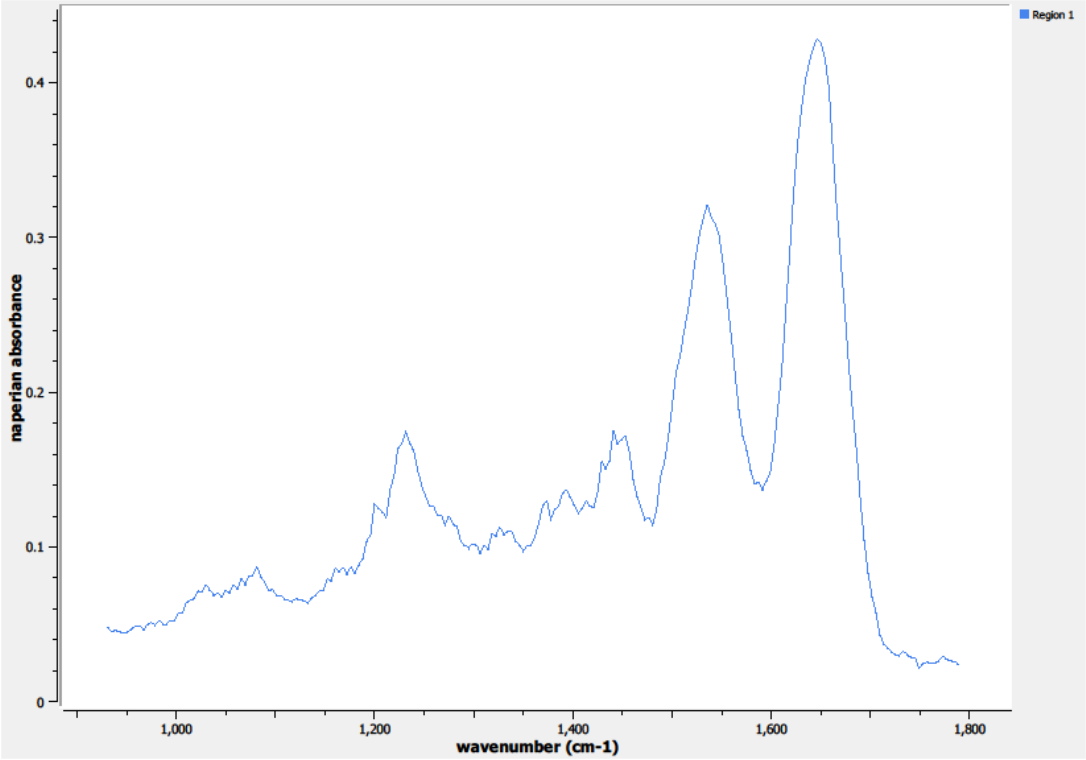



**Figure 0.77. Point R Spectra**

## A12. 13-02-A117a

| | |
|---|---|
| Unique ID | 13-02-A117a |
| Age | 60 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

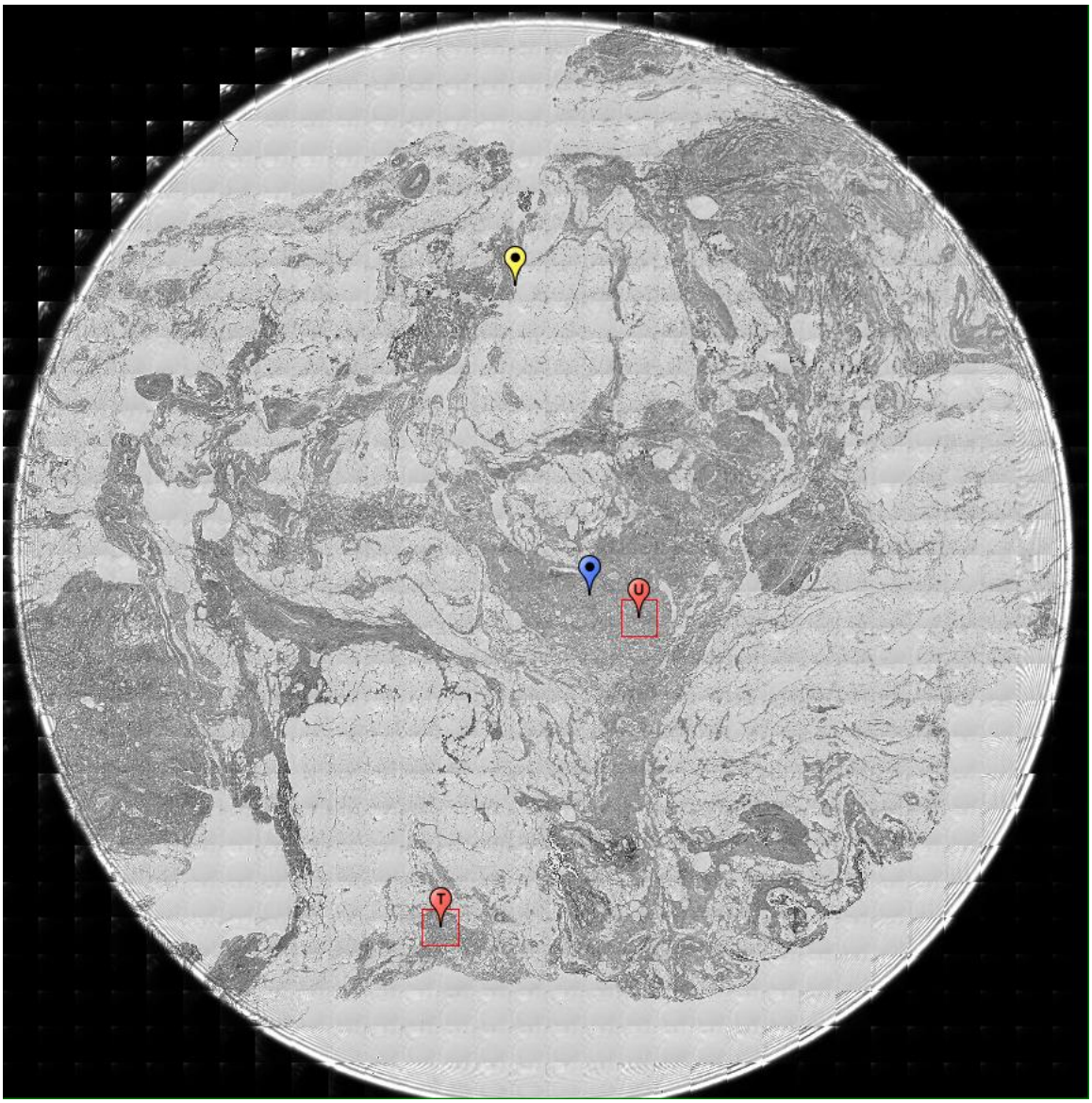



**Figure 0.78. Specimen 13-02-A117a, normal breast tissue.**

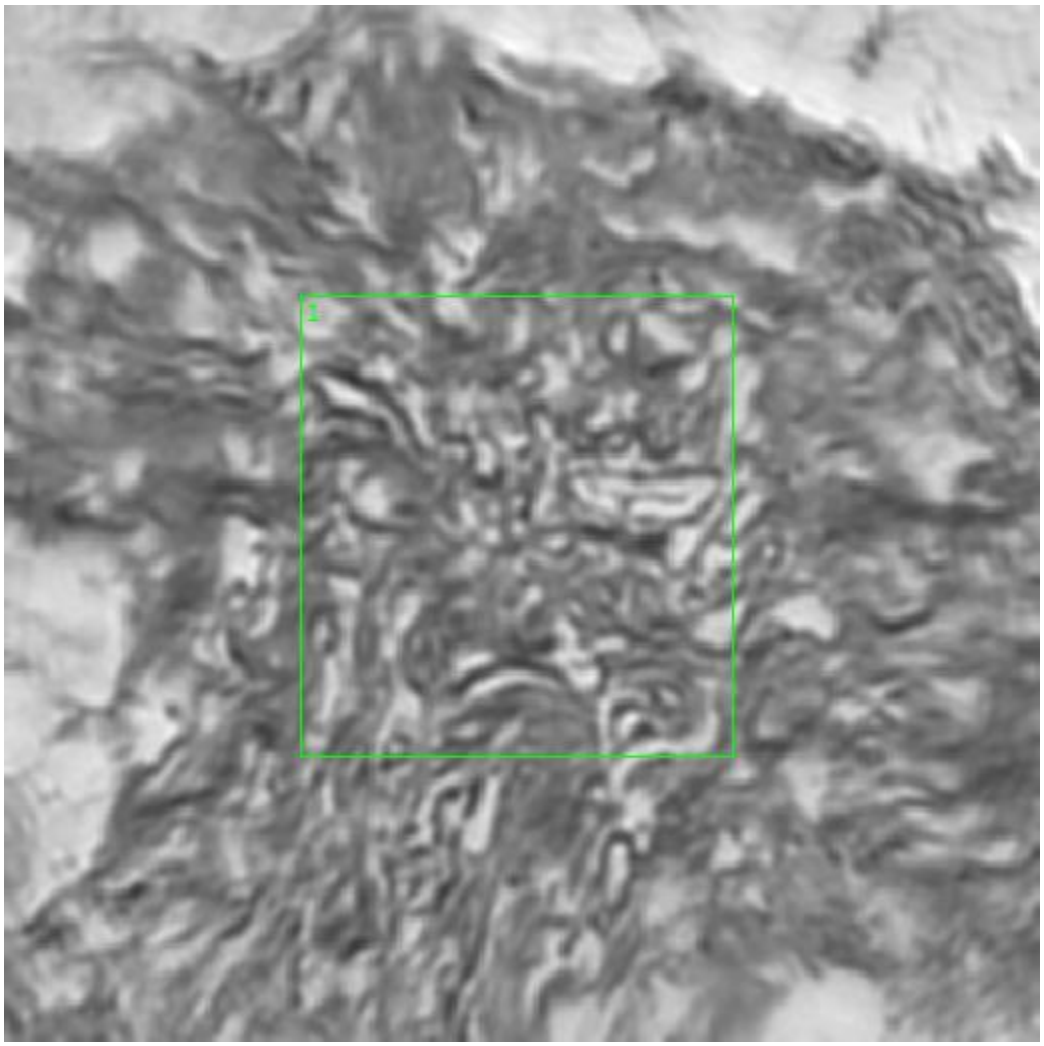


**Figure 0.79. Point T Regions of Interest**
**Region 1: breast**

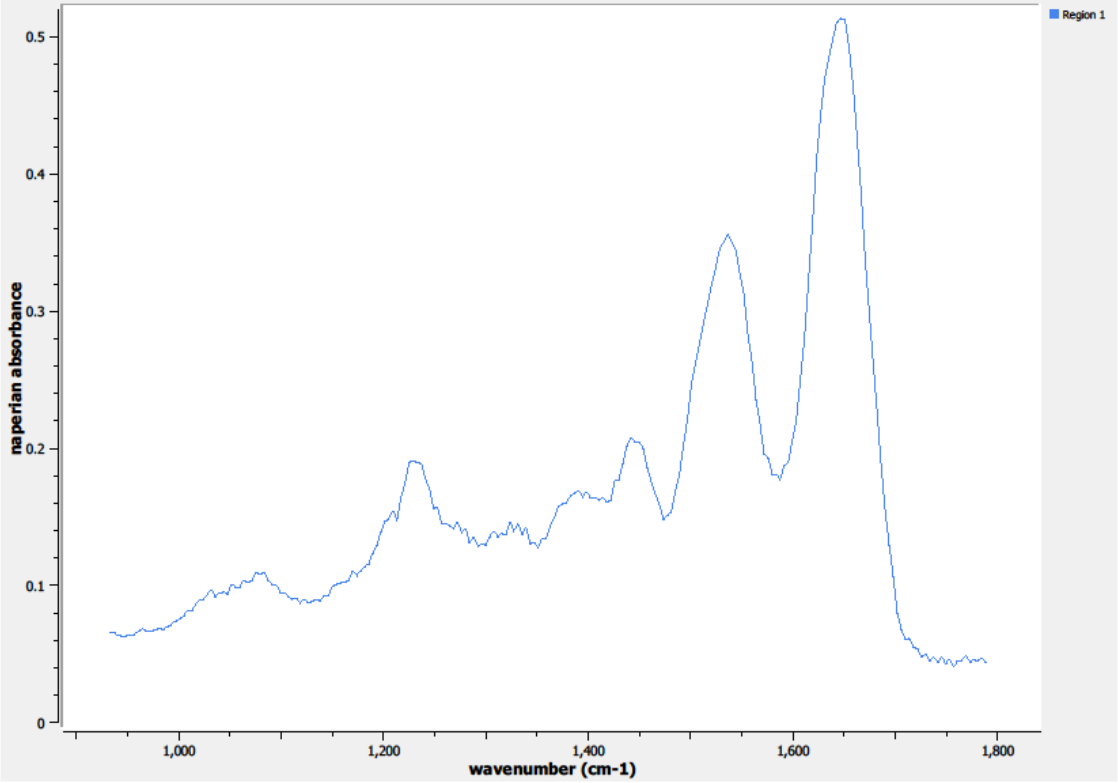



**Figure 0.80. Point T Spectra**

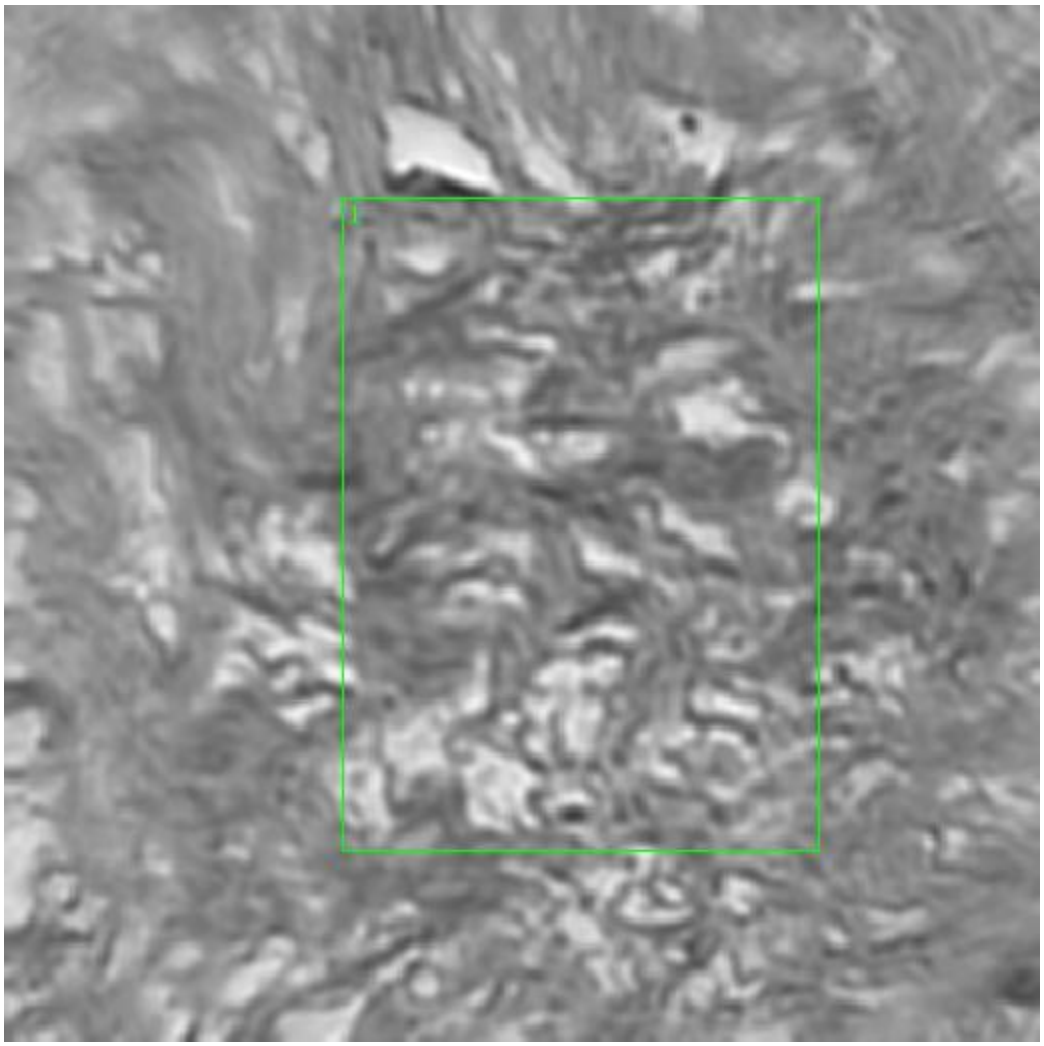


**Figure 0.81. Point U Regions of Interest**
**Region 1: breast**

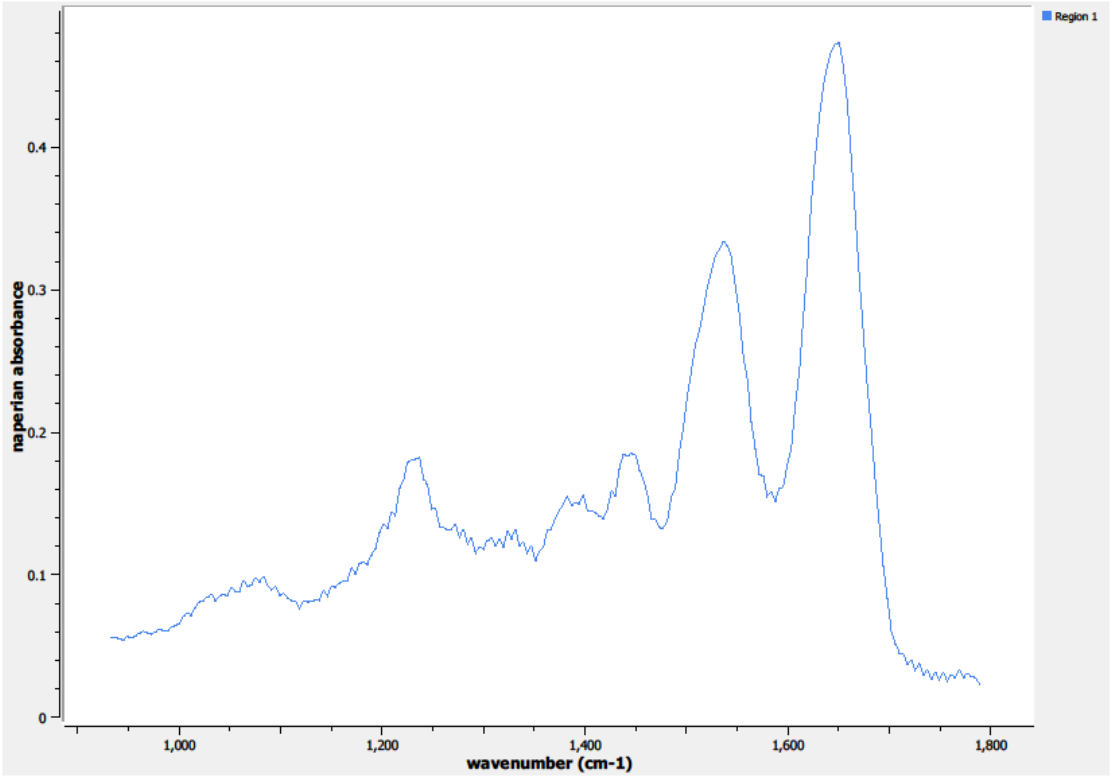



**Figure 0.82. Point U Spectra**

## A13. 13-04-A173a

| | |
|---|---|
| Unique ID | 13-04-A173a |
| Age | 63 |
| Race | black |
| Location of Sample Collection | unknown |
| Notes of Interest | Region Z too dense, may not use |

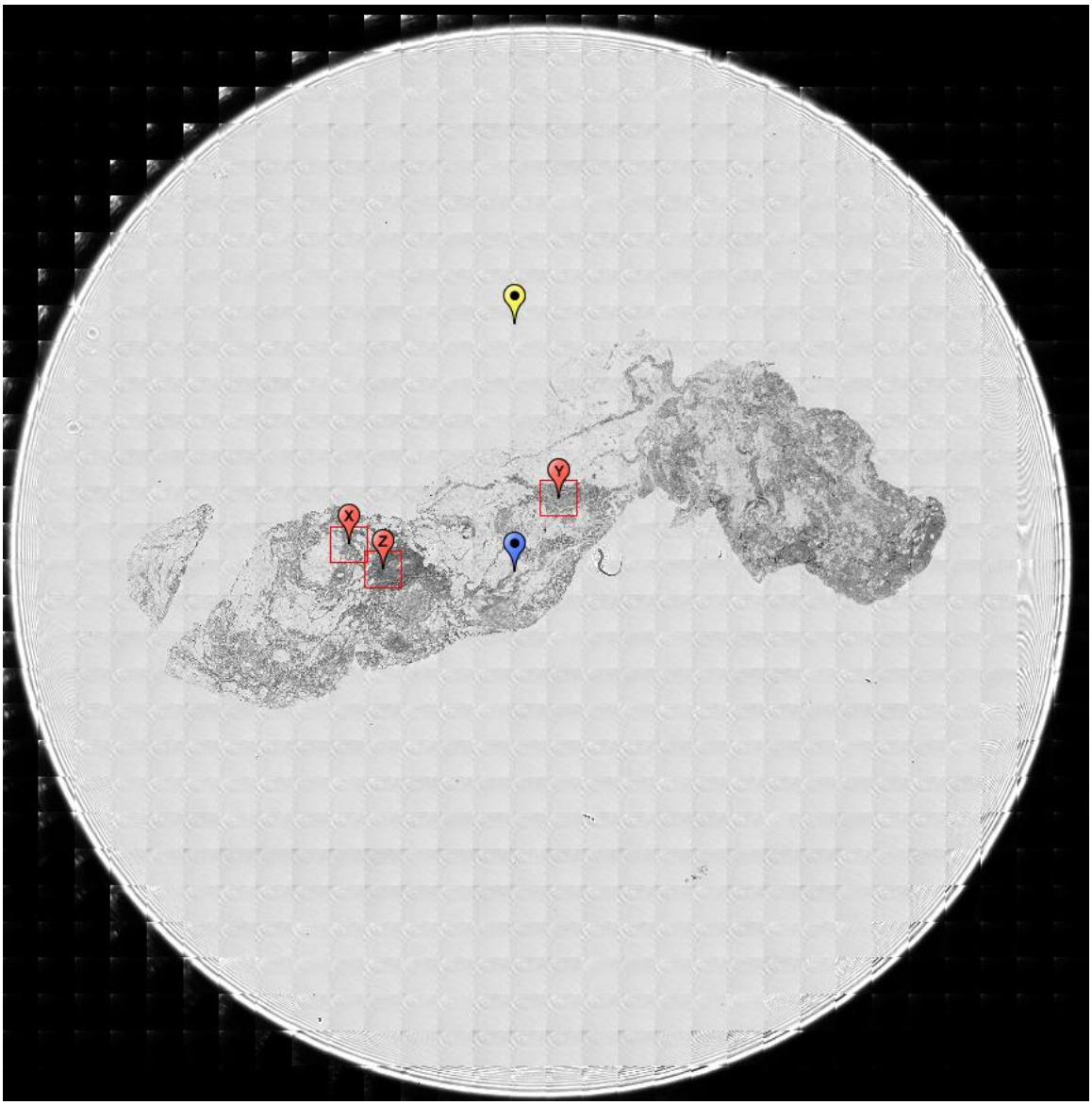



**Figure 0.83. Specimen 13-04-A173a, normal breast tissue.**

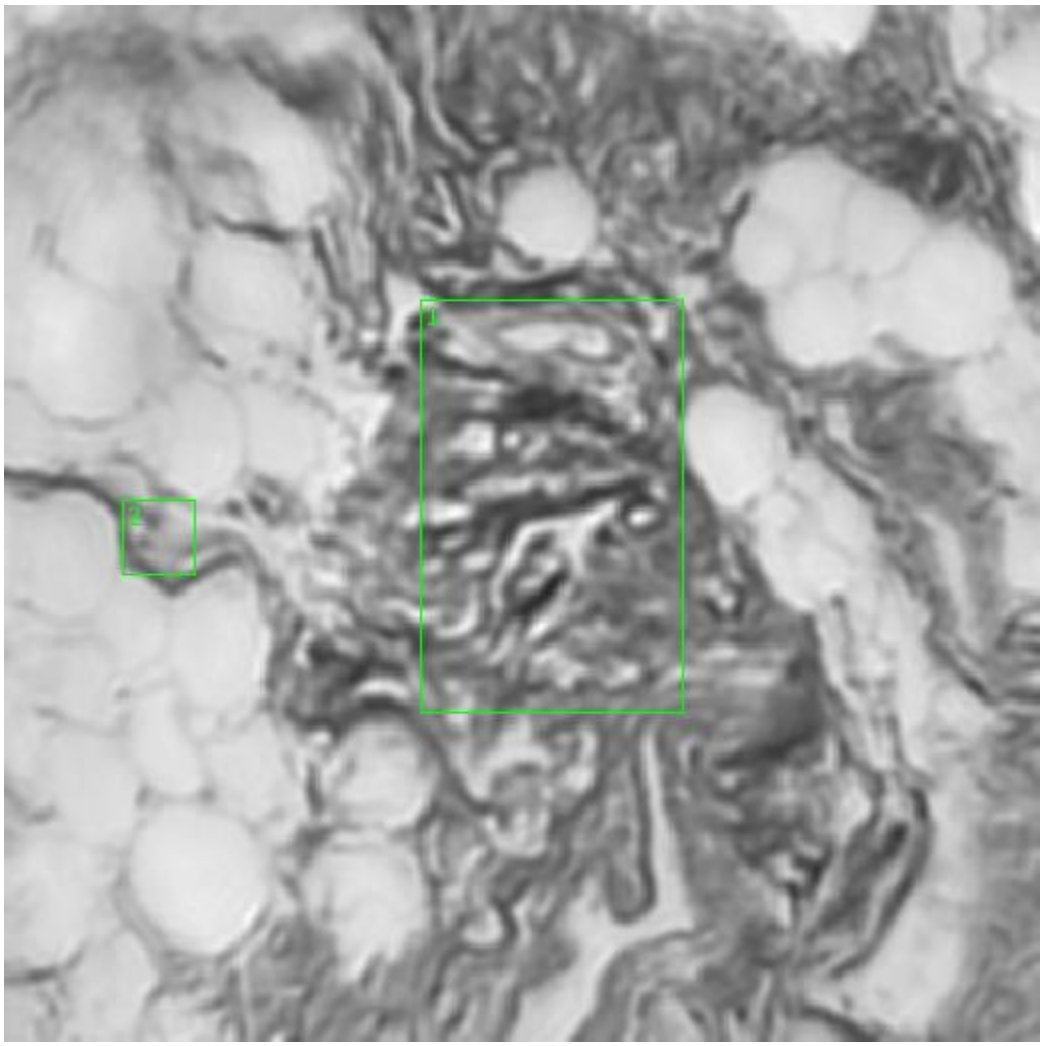


**Figure 0.84. Point X Regions of Interest**
**Region 1: breast**
**Region 2: concentrated region, breast**

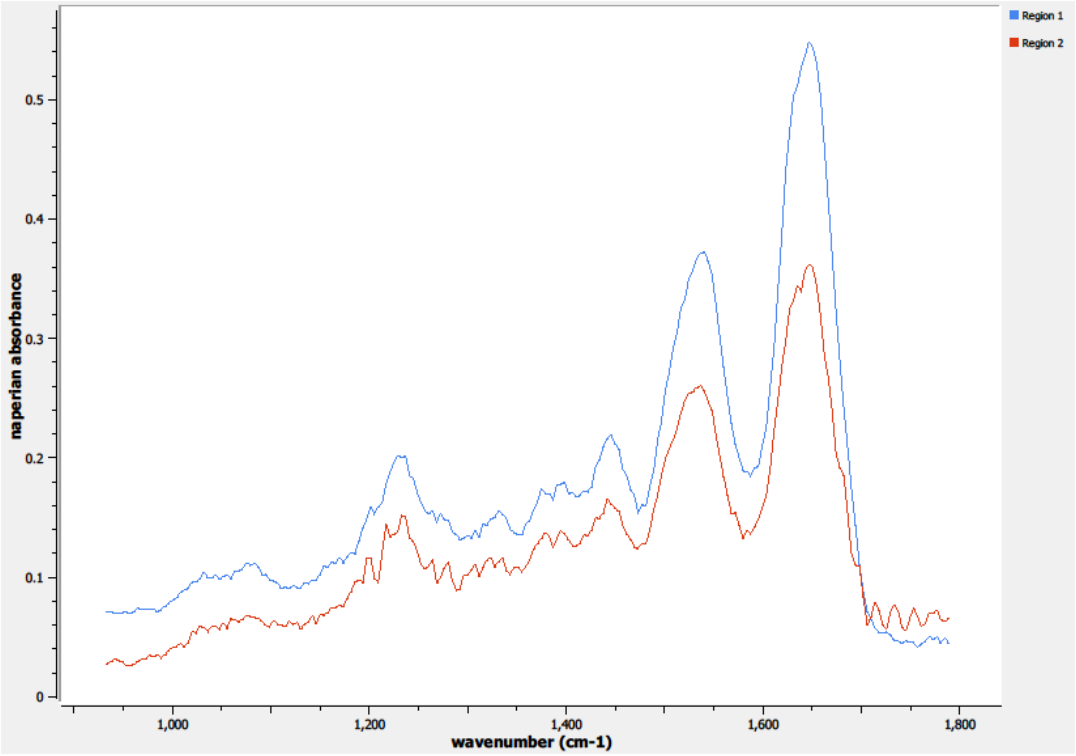



**Figure 0.85. Point X Spectra**

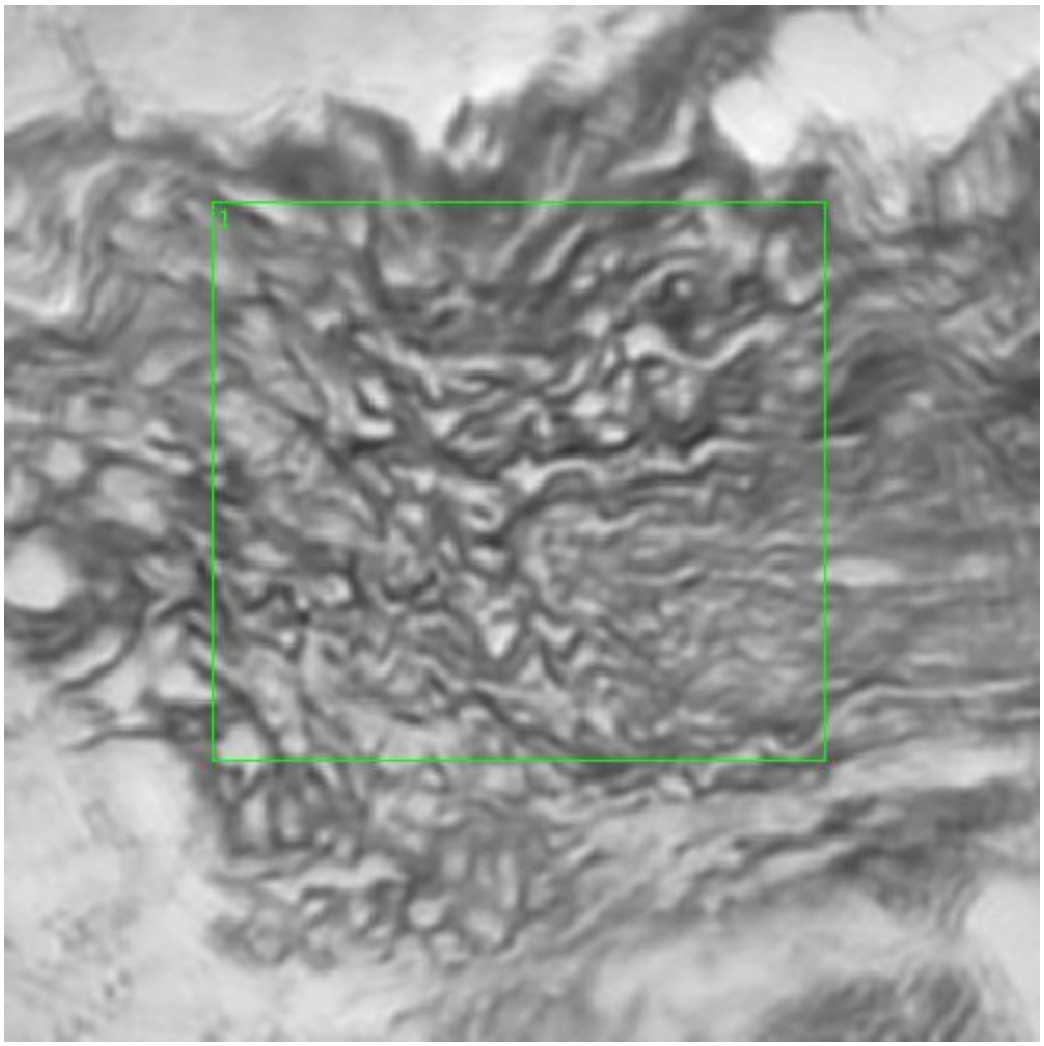



**Figure 0.86. Point Y Regions of Interest**
**Region 1: breast**

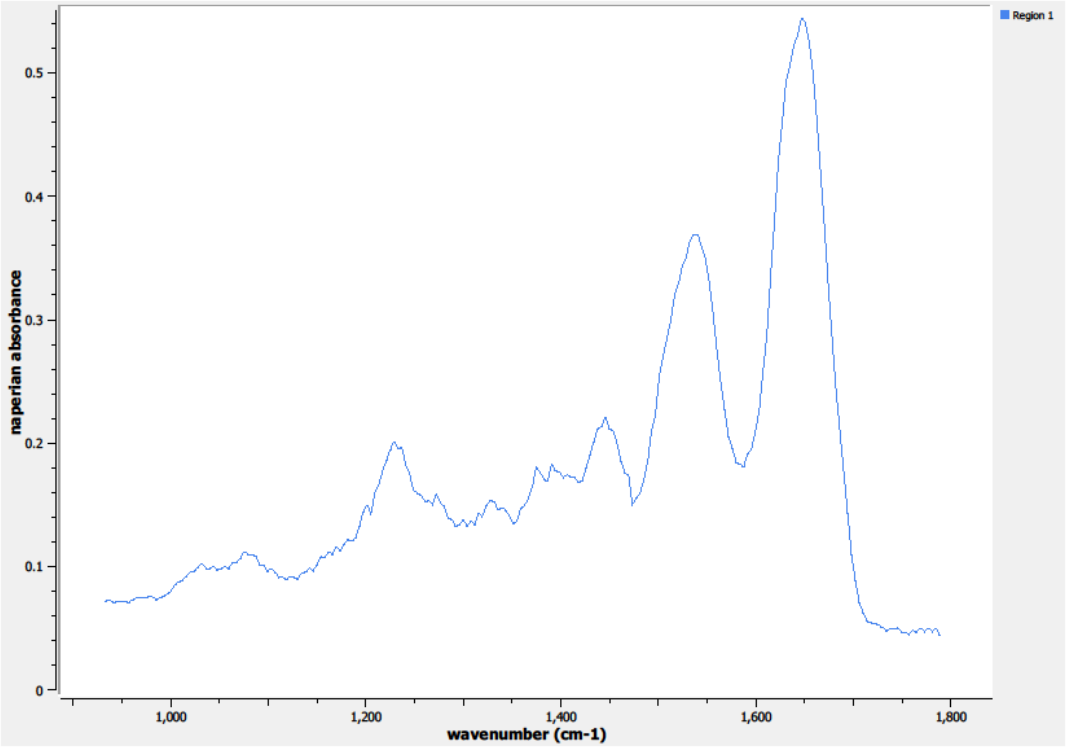



**Figure 0.87. Point Y Spectra**

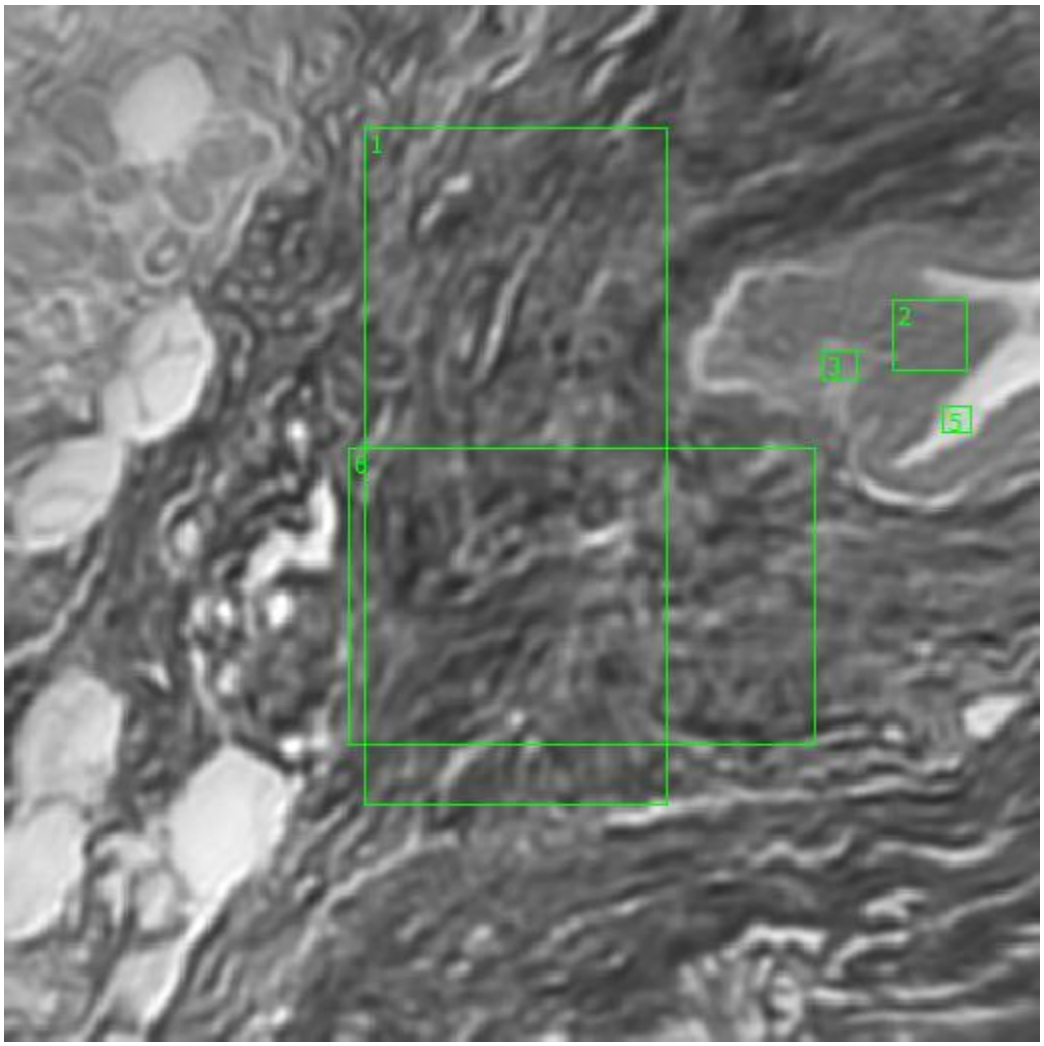



**Figure 0.88. Point Z Regions of Interest**
**Regions 1, 6: normal breast tissue**
**Regions 2, 3: membrane surrounding glandular structure**
**Region 5: interior glandular structure**

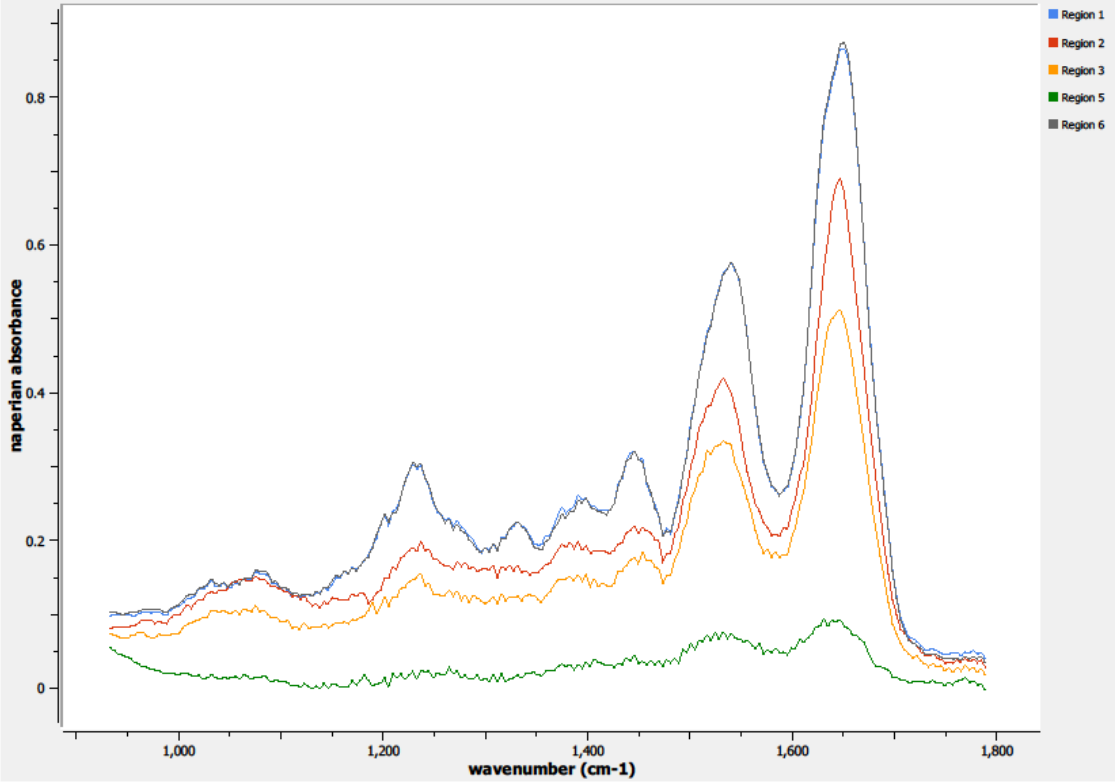



**Figure 0.89. Point Z Spectra**

## A14. 13-04-A184a

| | |
|---|---|
| Unique ID | 13-04-A184a |
| Age | 62 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

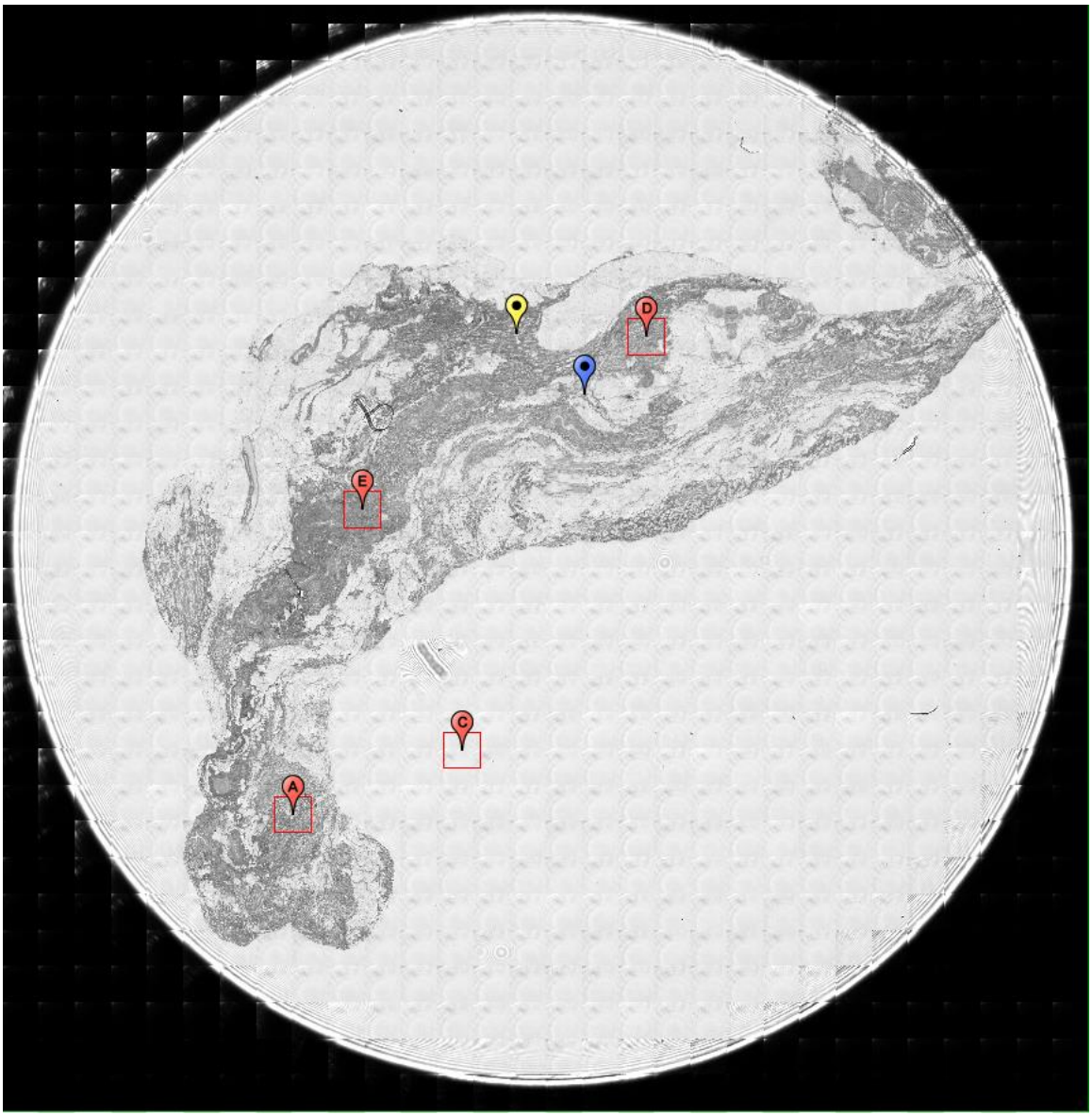



**Figure 0.90. Specimen 13-04-A184a, normal breast tissue.**

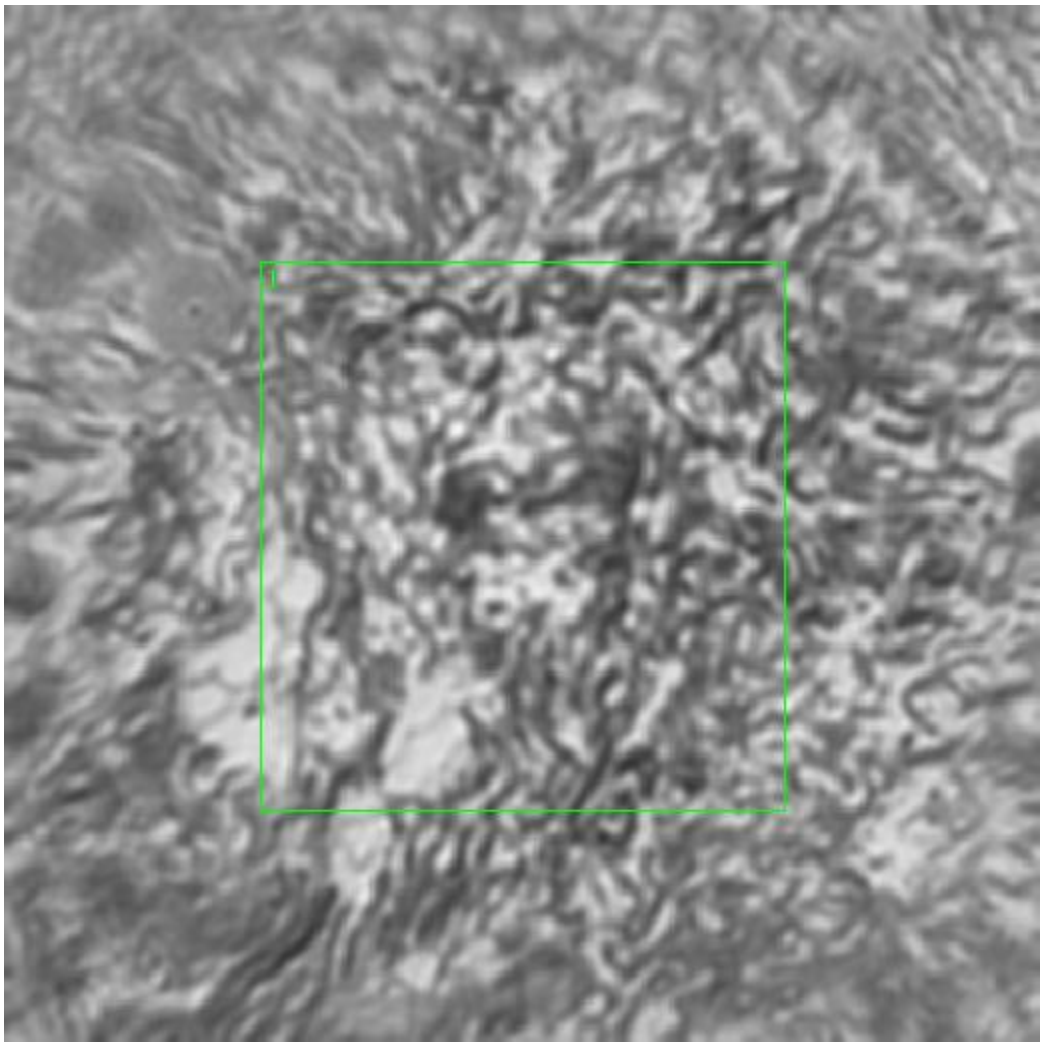


**Figure 0.91. Point A Regions of Interest**
**Region 1: breast**

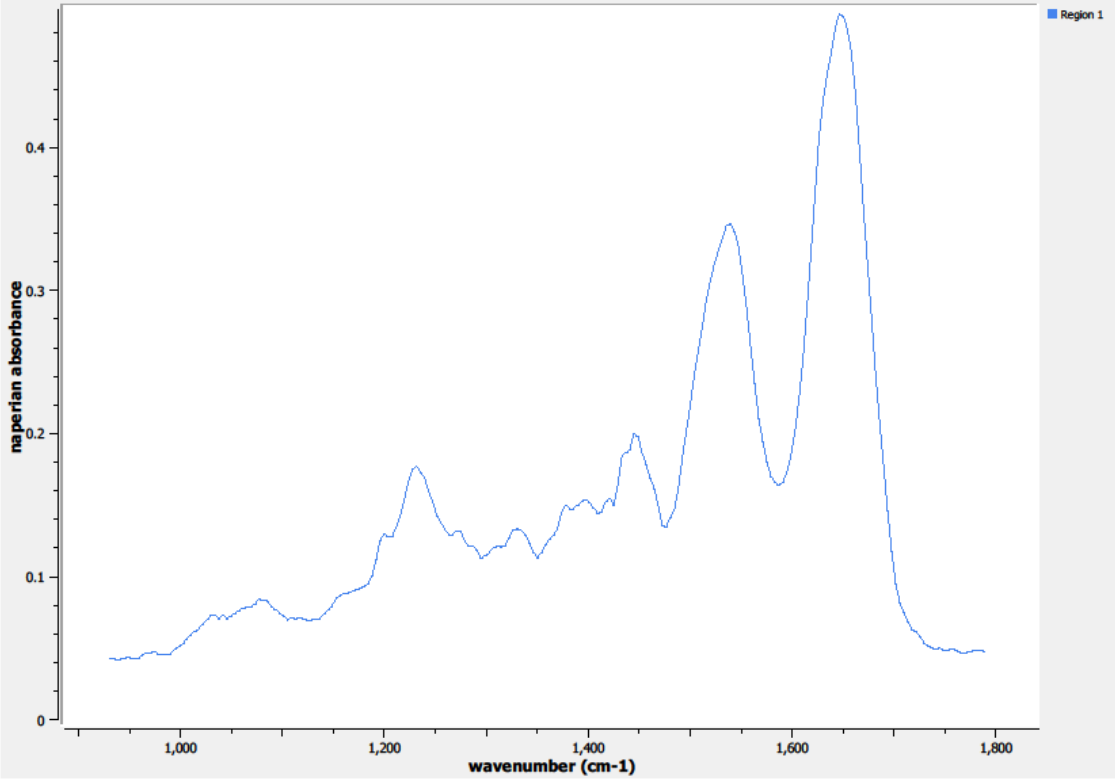



**Figure 0.92. Point A Spectra**

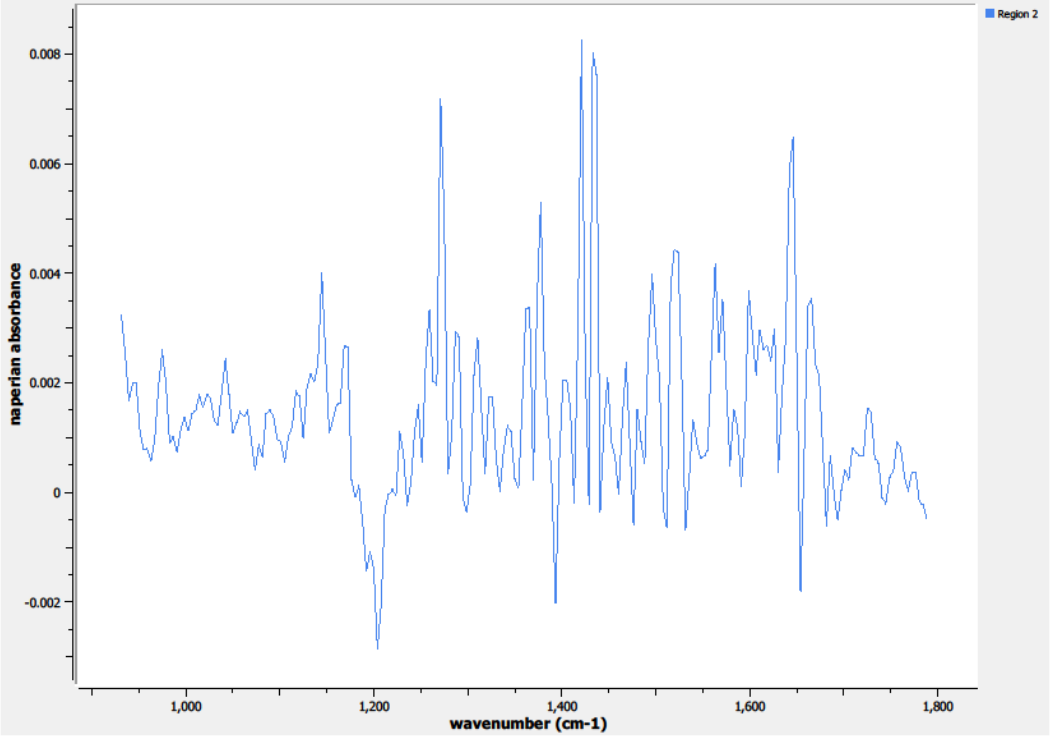



**Figure 0.93. Point C Background Spectra**

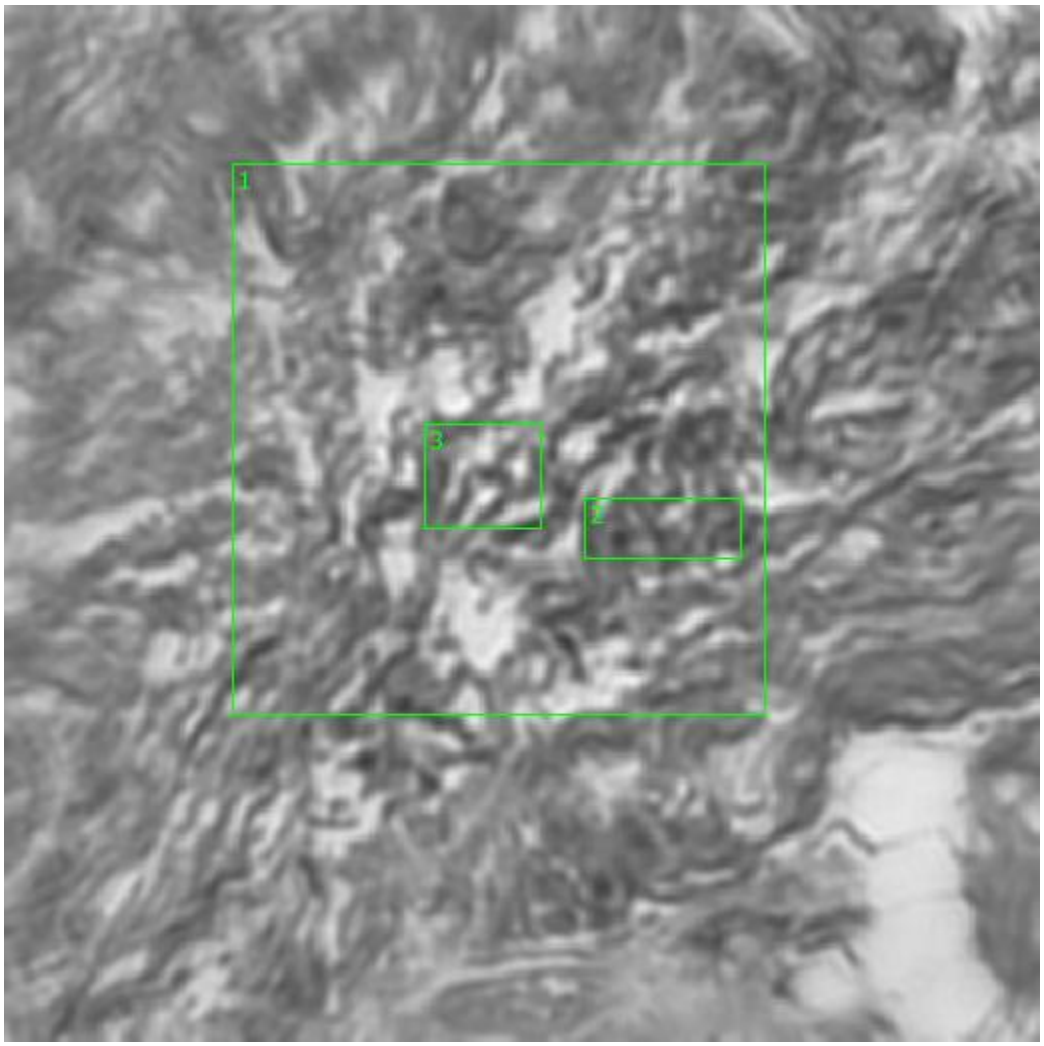



**Figure 0.94. Point D Regions of Interest**
**Region 1: breast**
**Region 2, 3: concentrated regions, breast**

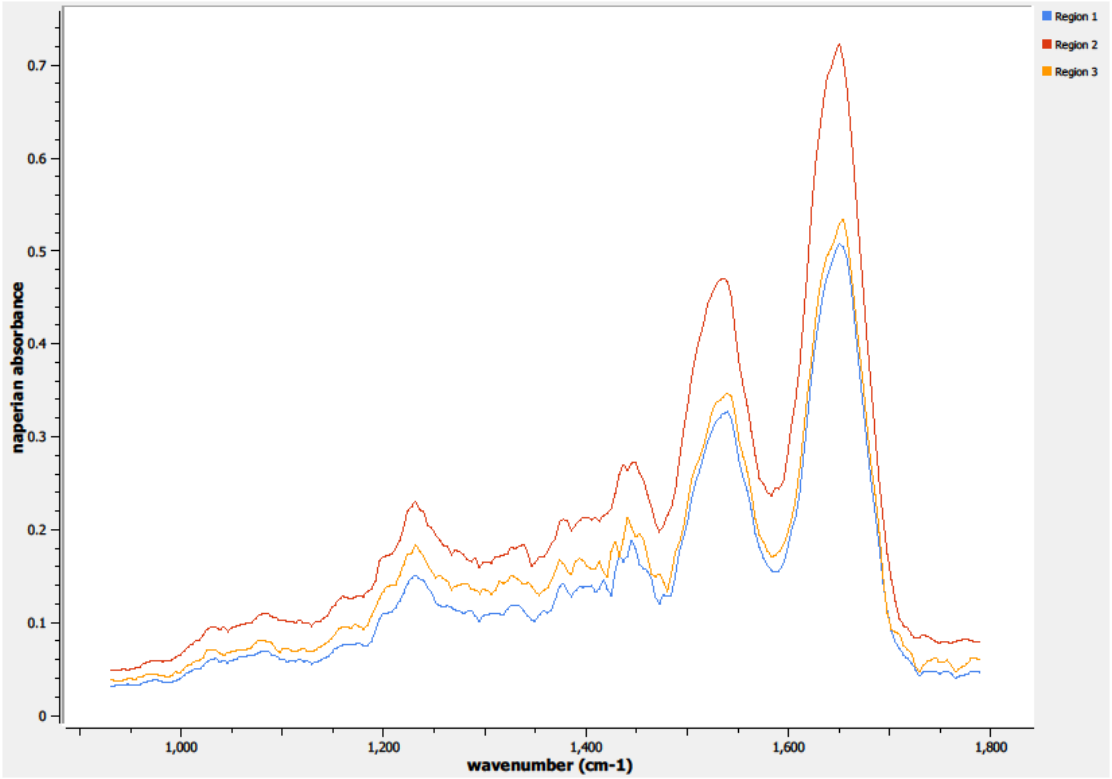



**Figure 0.95. Point D Spectra**

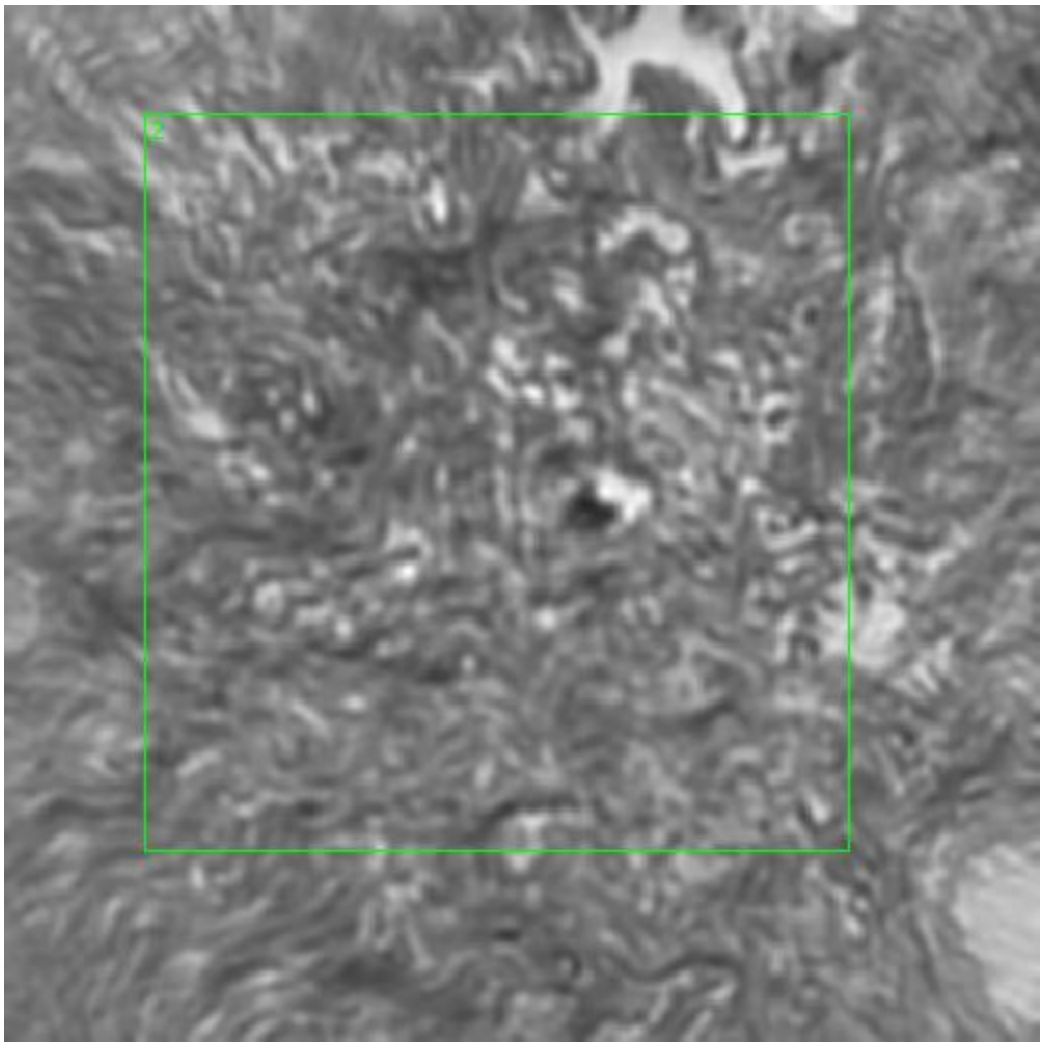


**Figure 0.96. Point E Regions of Interest**
**Region 1: breast**

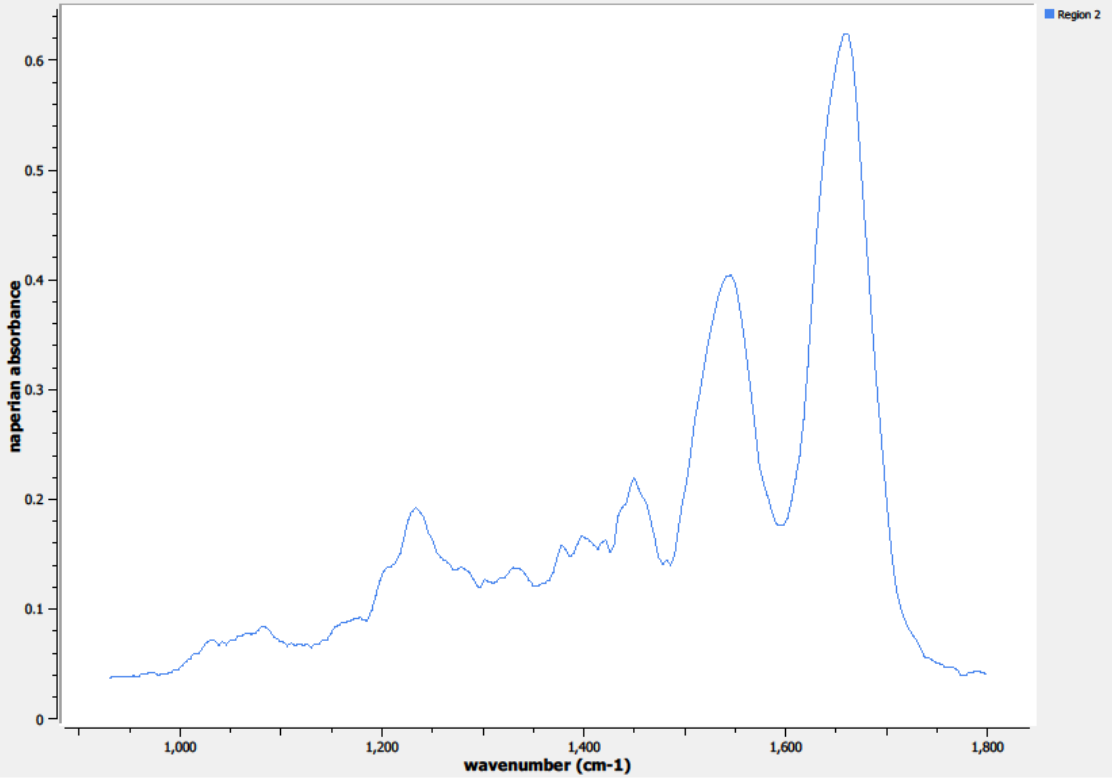



**Figure 0.97. Point E Spectra**

## A15. 13-04-A204a

| | |
|---|---|
| Unique ID | 13-04-A204a |
| Age | 60 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

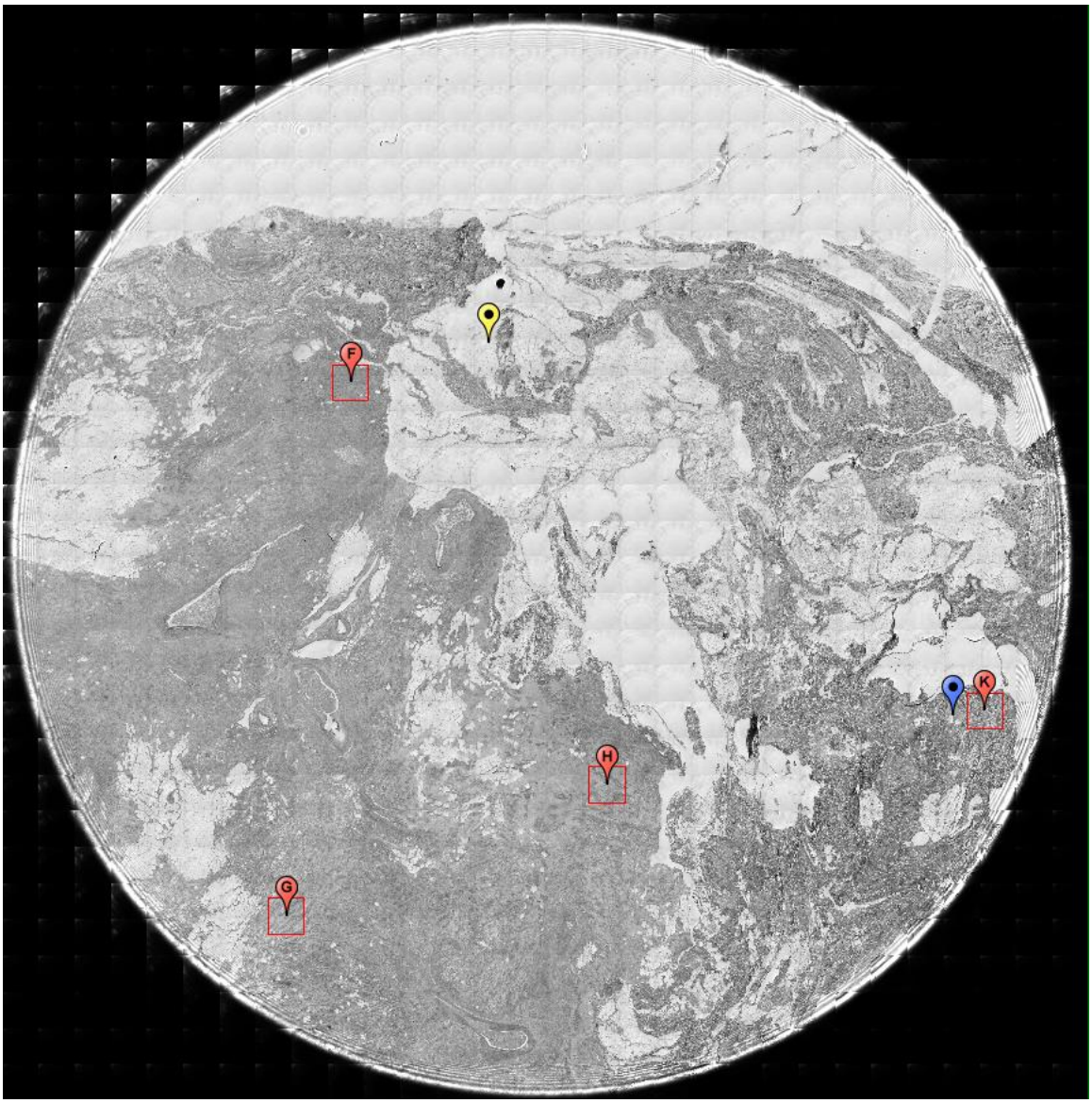



**Figure 0.98. Specimen 13-04-A204a, normal breast tissue.**

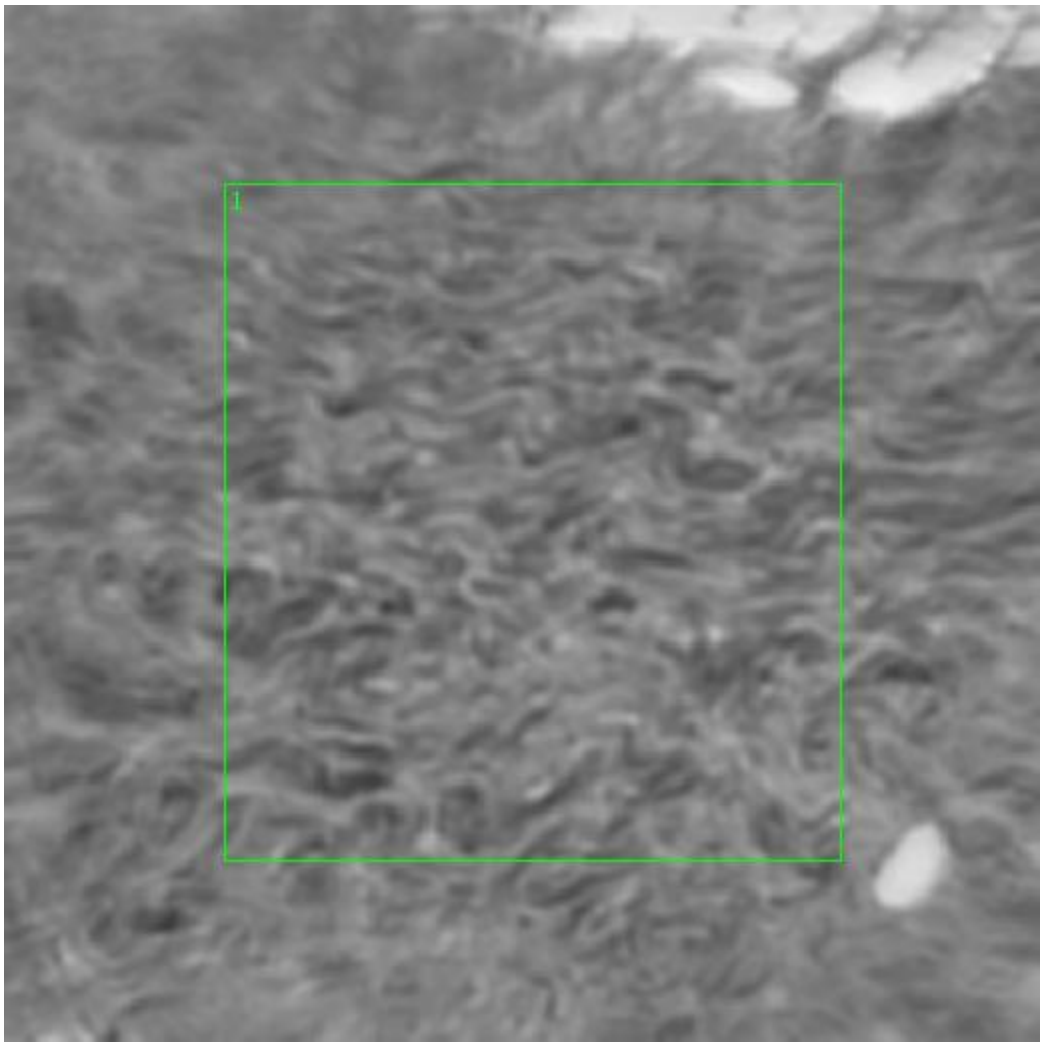



**Figure 0.99. Point F Regions of Interest**
**Region 1: breast**

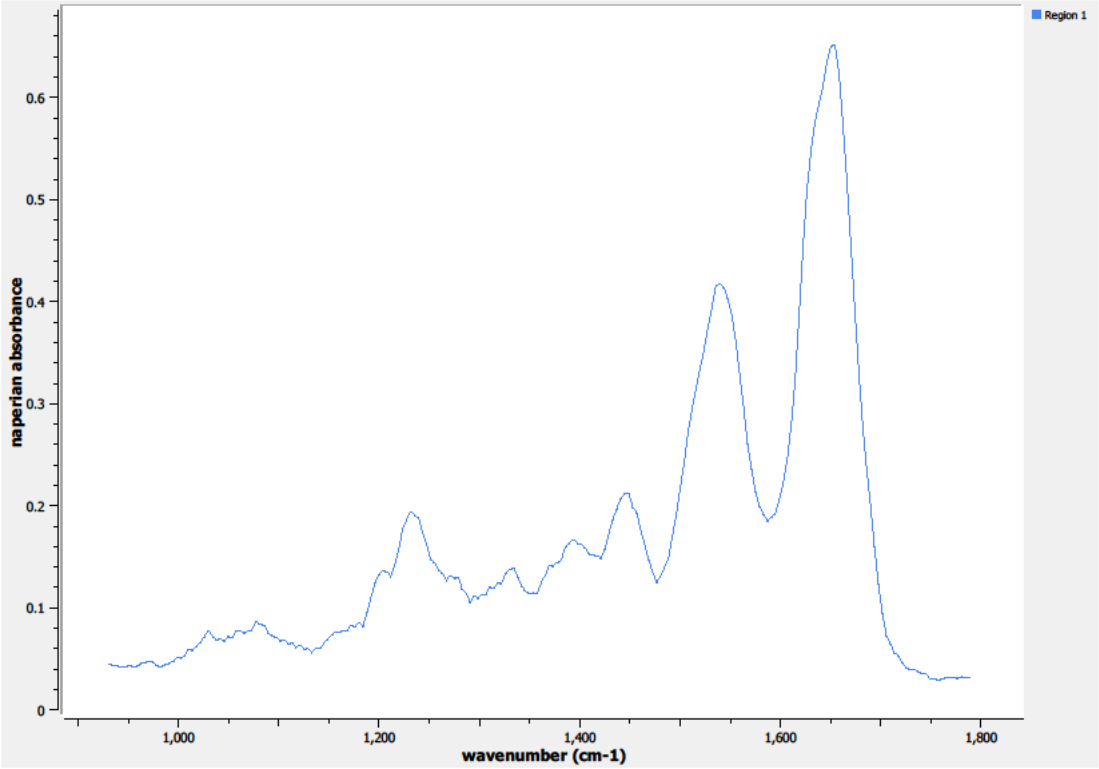



**Figure 0.100. Point F Spectra**

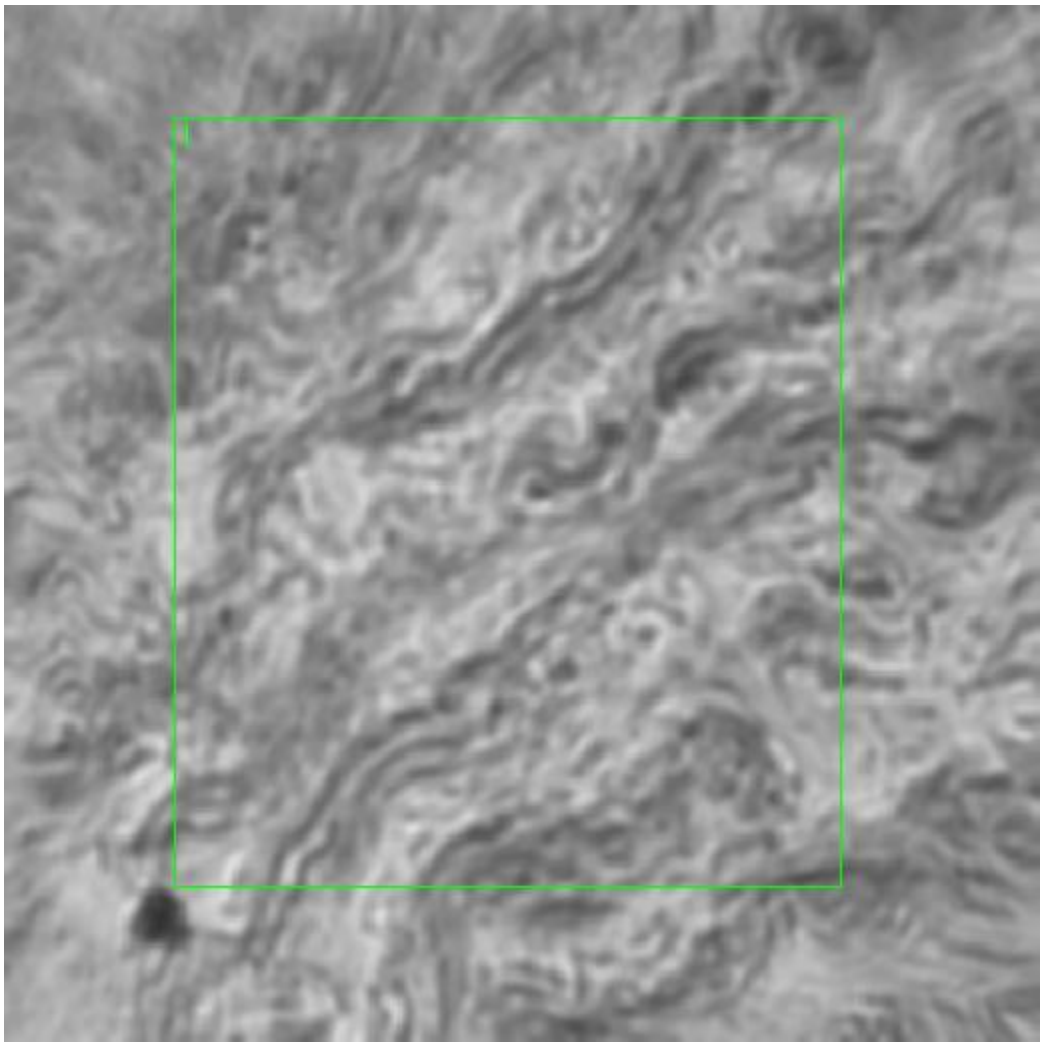


**Figure 0.101. Point G Regions of Interest**
**Region 1: breast**

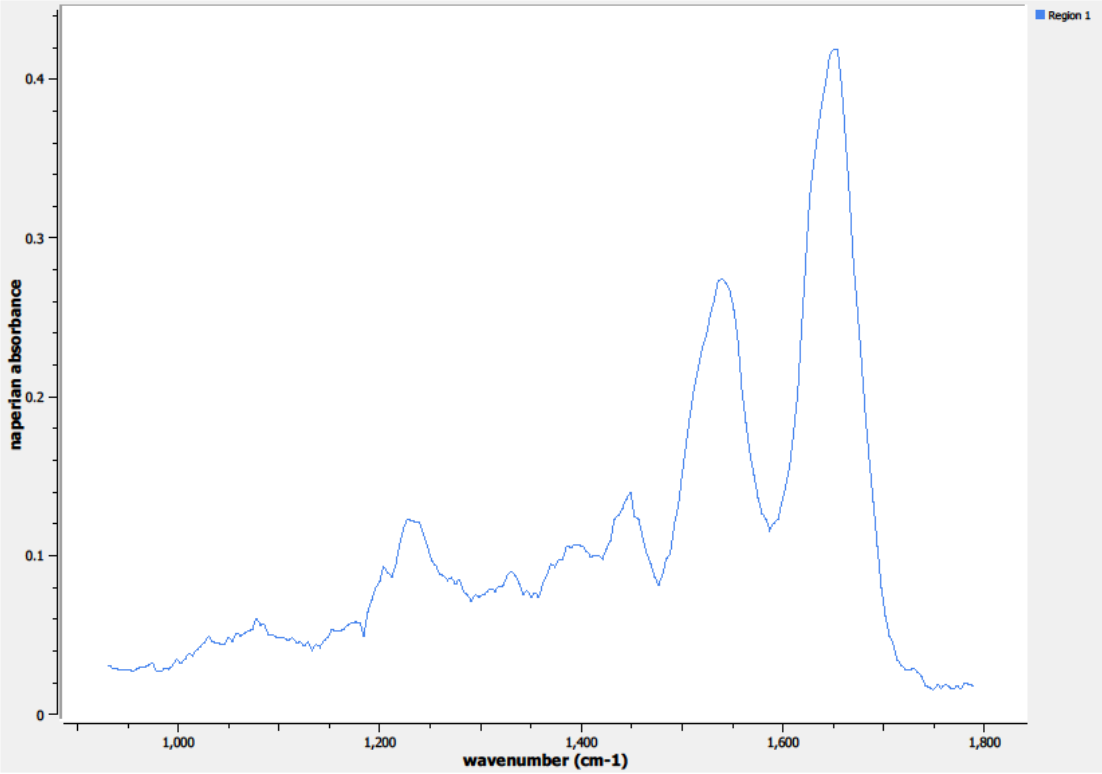



**Figure 0.102. Point G Spectra**

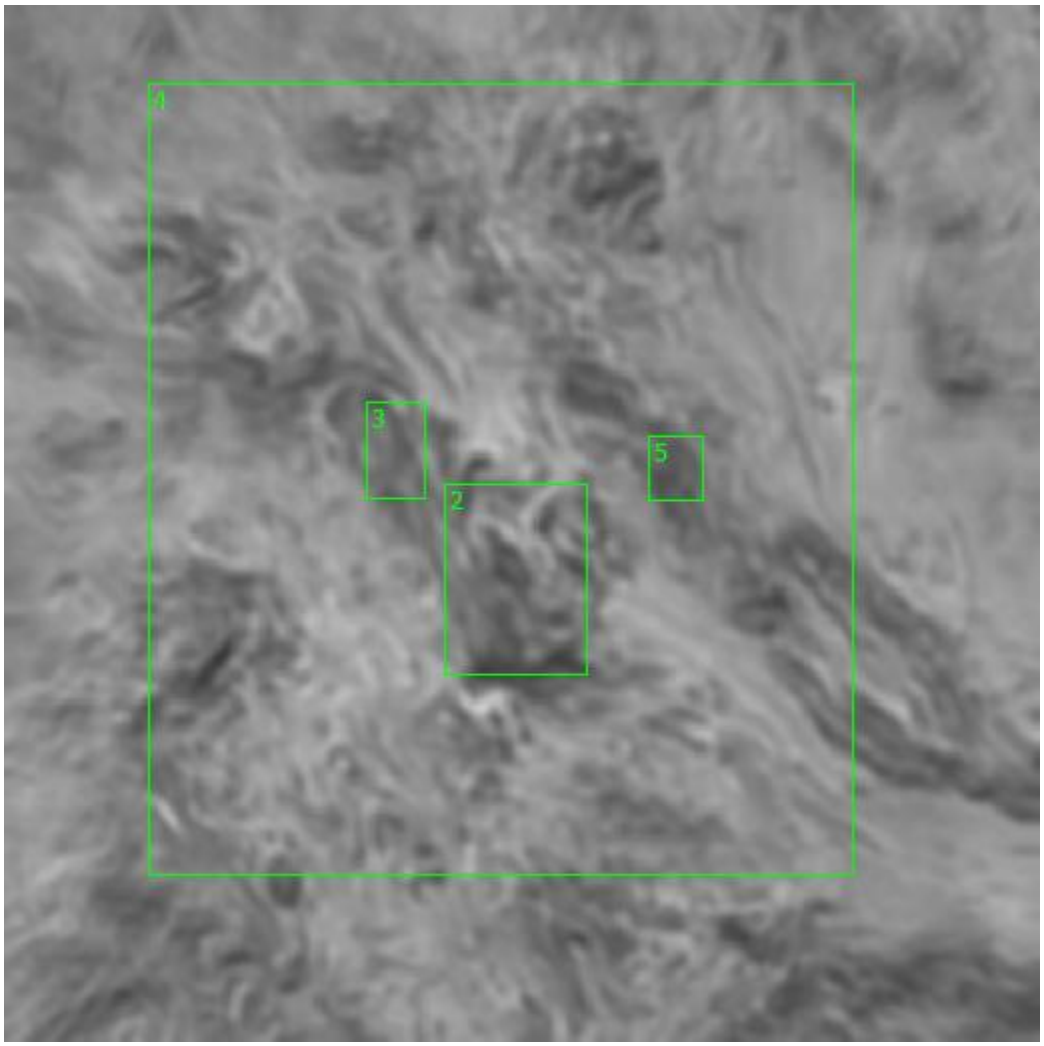



**Figure 0.103. Point H Regions of Interest**
**Regions 2, 3, 5: concentrated regions, breast**
**Region 4: breast**

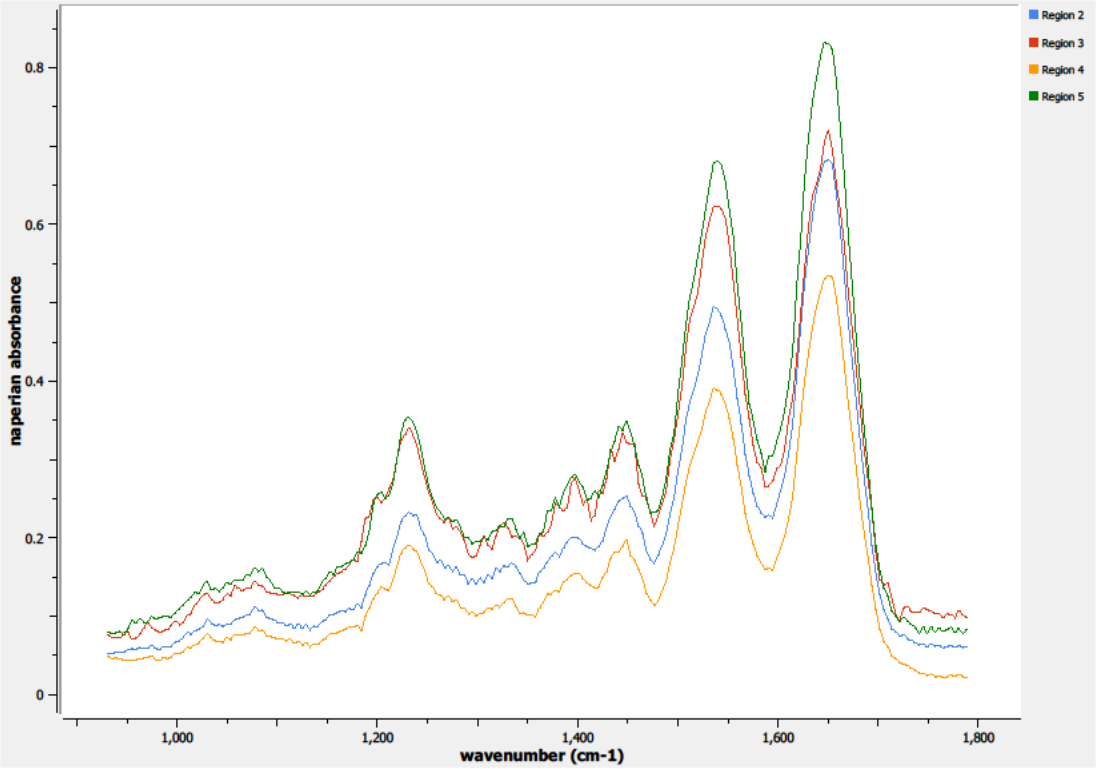



**Figure 0.104. Point H Spectra**

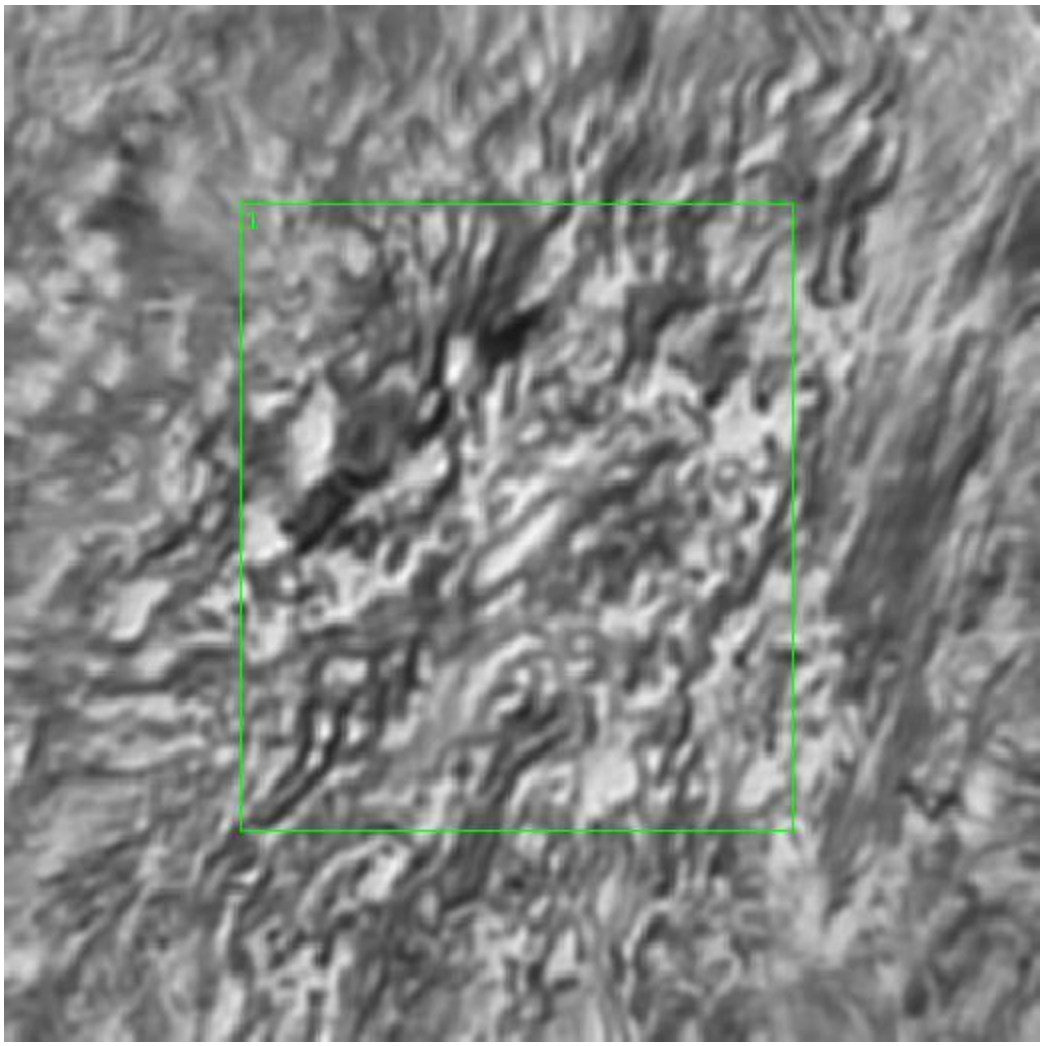


**Figure 0.105. Point K Regions of Interest**
**Region 1: breast**

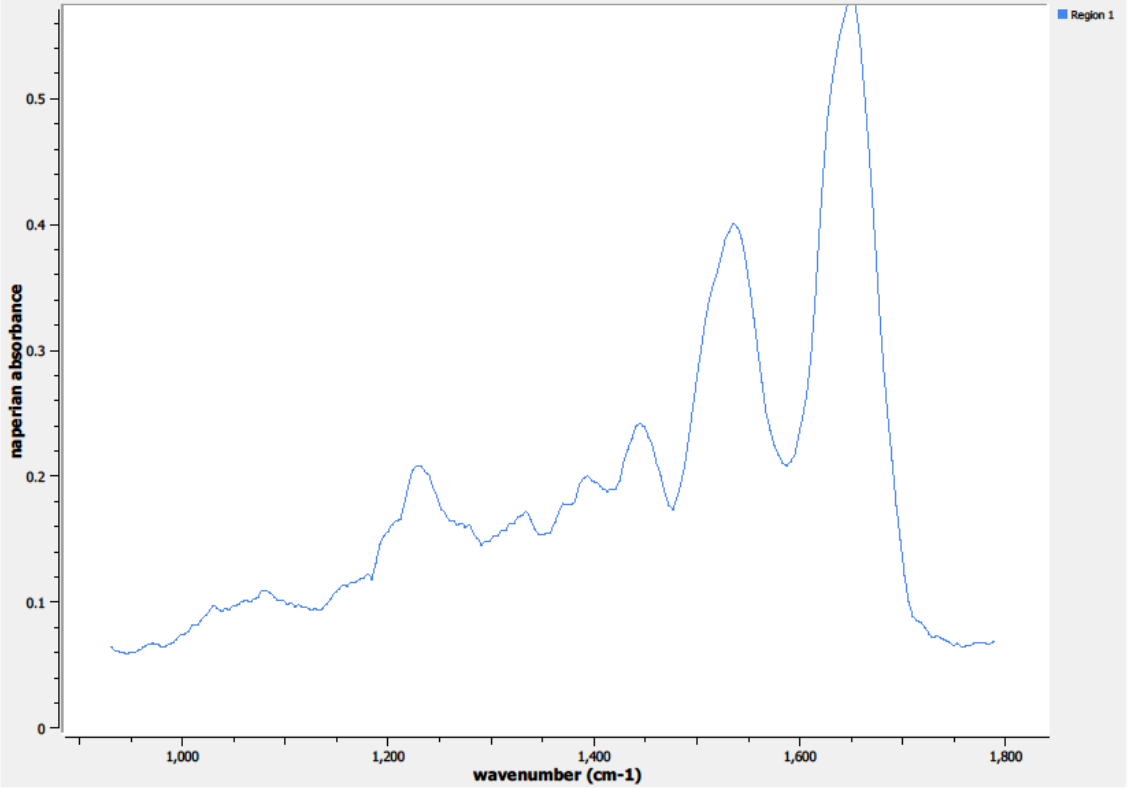



**Figure 0.106. Point K Spectra**

## A16. 13-05-A012a

| | |
|---|---|
| Unique ID | 13-05-A012a |
| Age | 61 |
| Race | white |
| Location of Sample Collection | unknown |
| Notes of Interest | unable to obtain clean background, unused in data results |

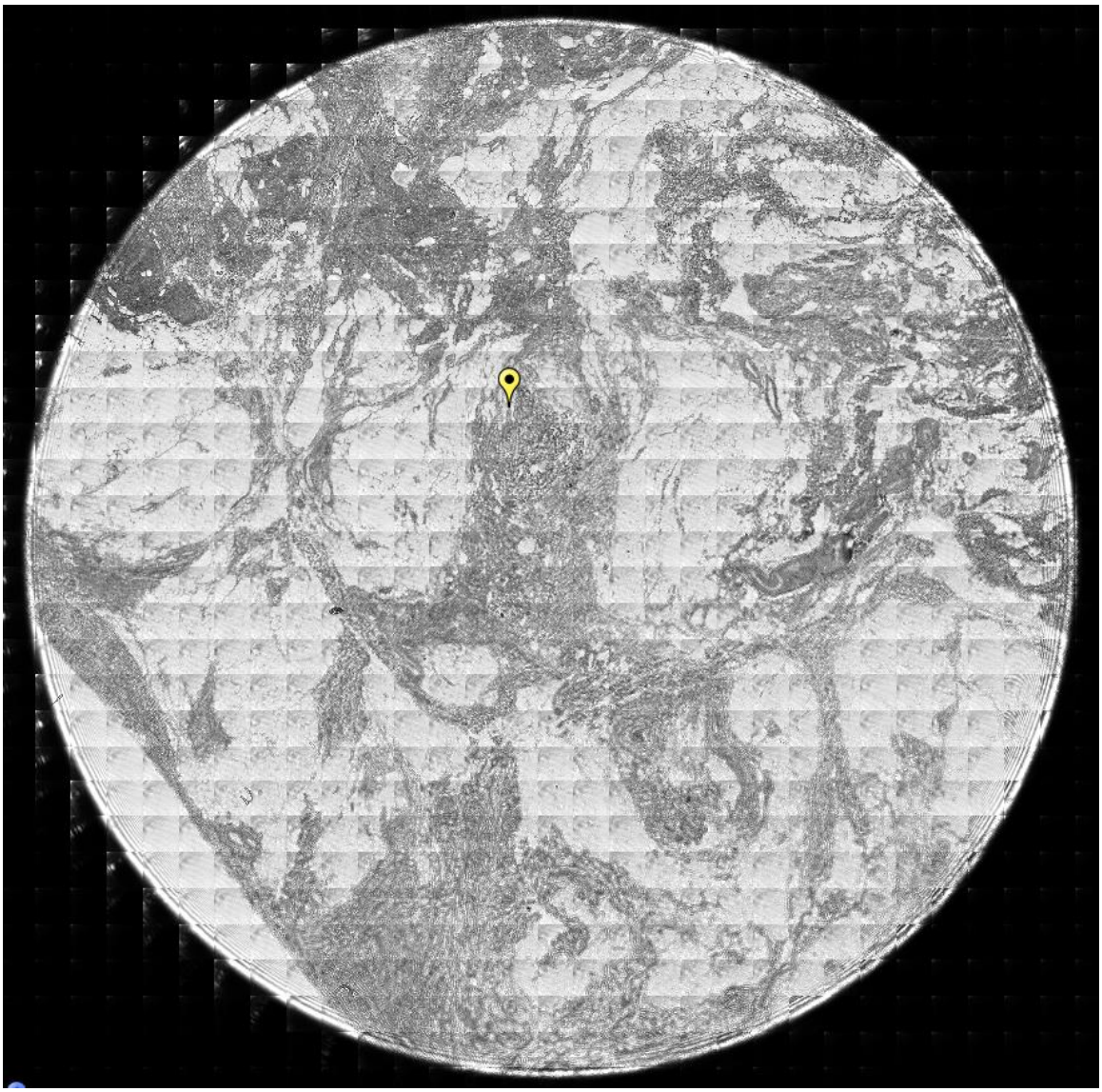


**Figure 0.107. Specimen 13-05-A012a, normal breast tissue.**

## A17. 13-05-A113a

| | |
|---|---|
| Unique ID | 13-05-A113a |
| Age | 60 |

| | |
|---|---|
| Race | black |
| Location of Sample Collection | unknown |
| Notes of Interest | none |

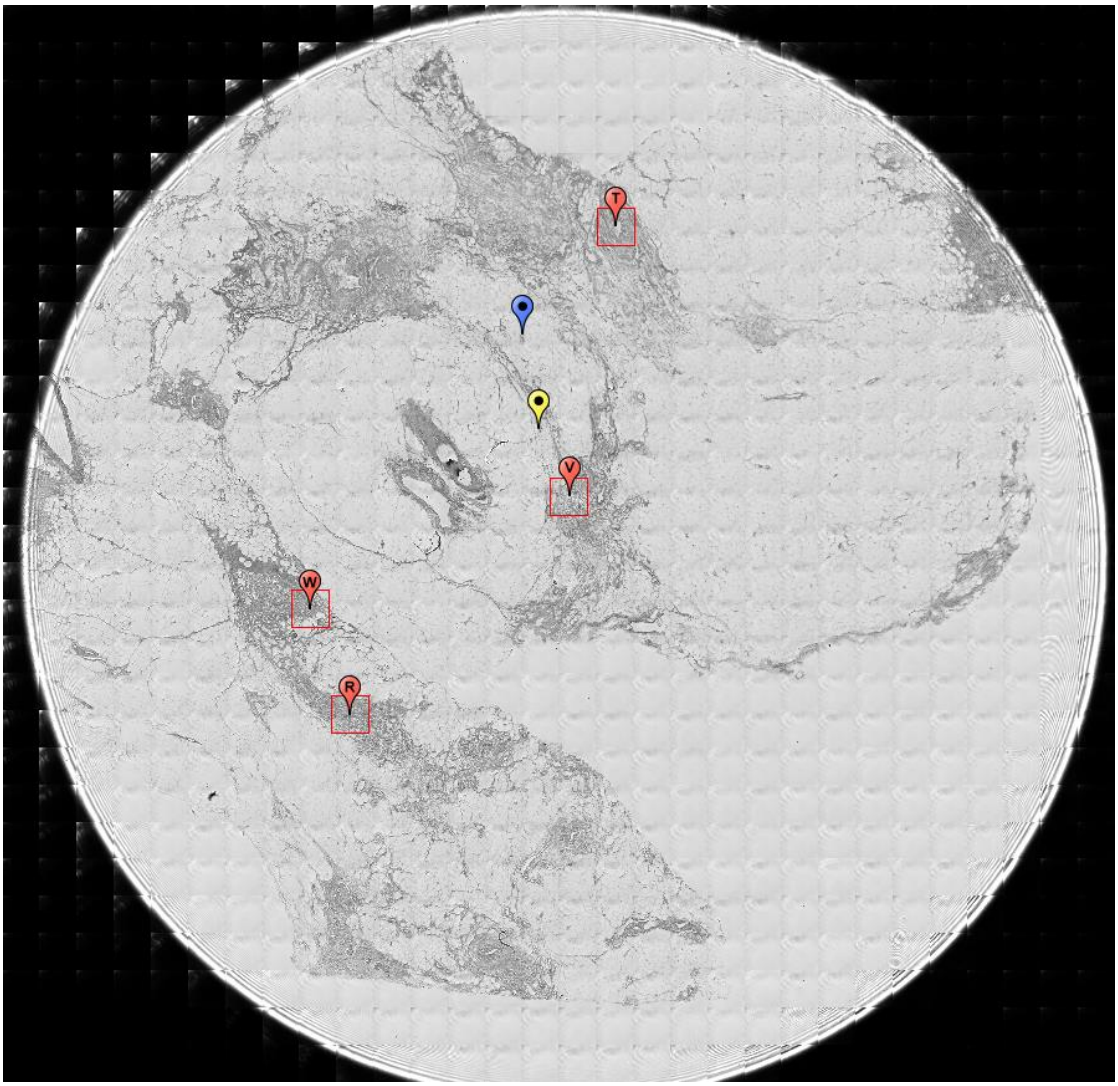



**Figure 0.108. Specimen 13-05-A113a, normal breast tissue.**

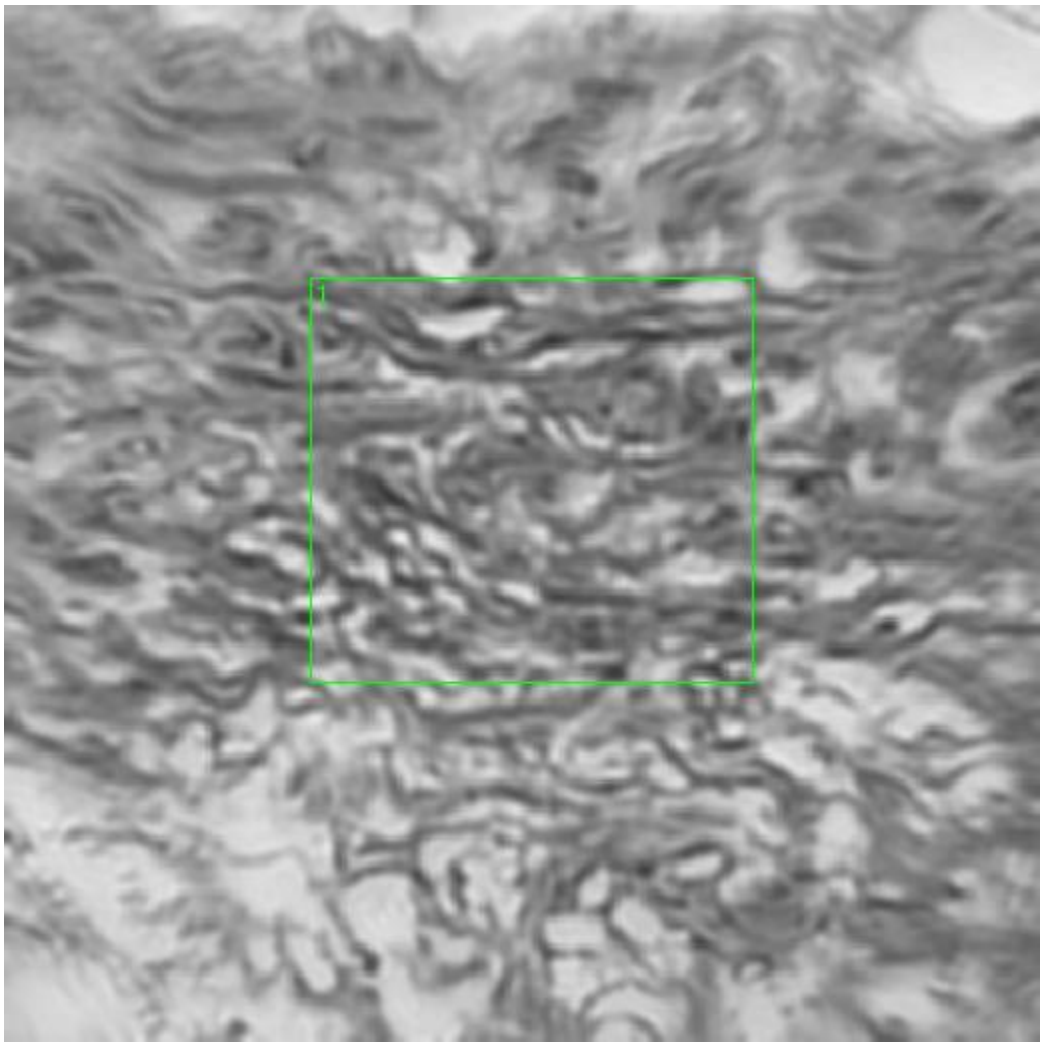


**Figure 0.109. Point R Regions of Interest**
**Region 1: breast**

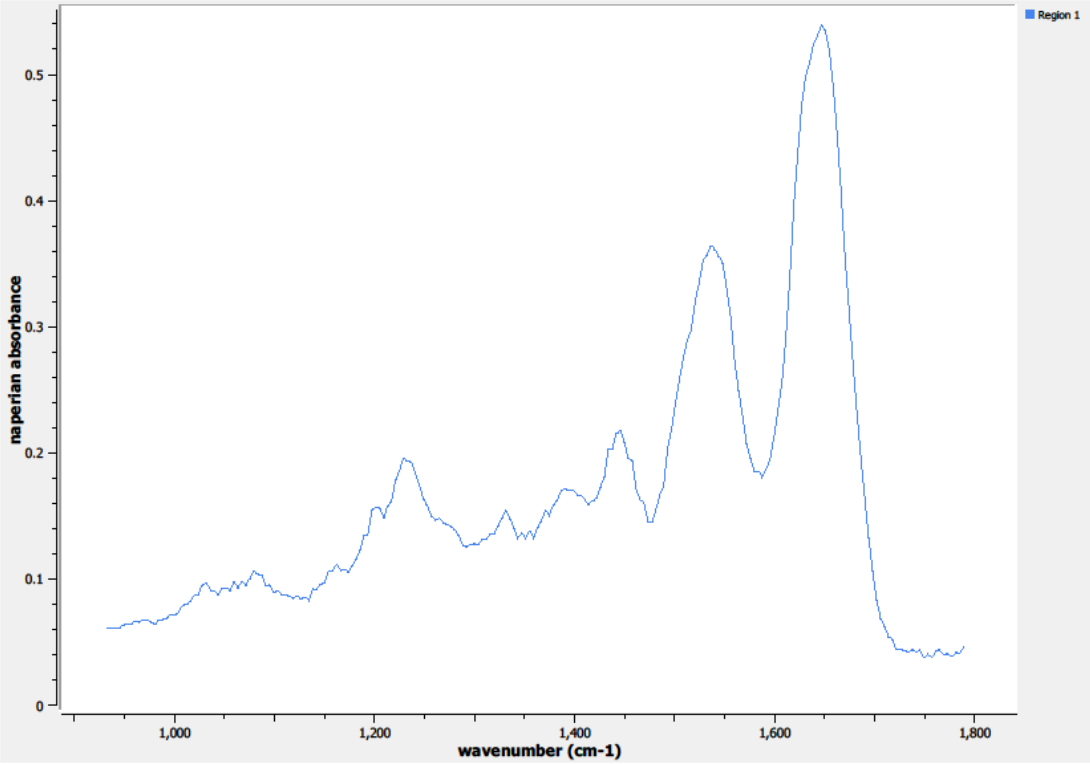



**Figure 0.110. Point R Spectra**

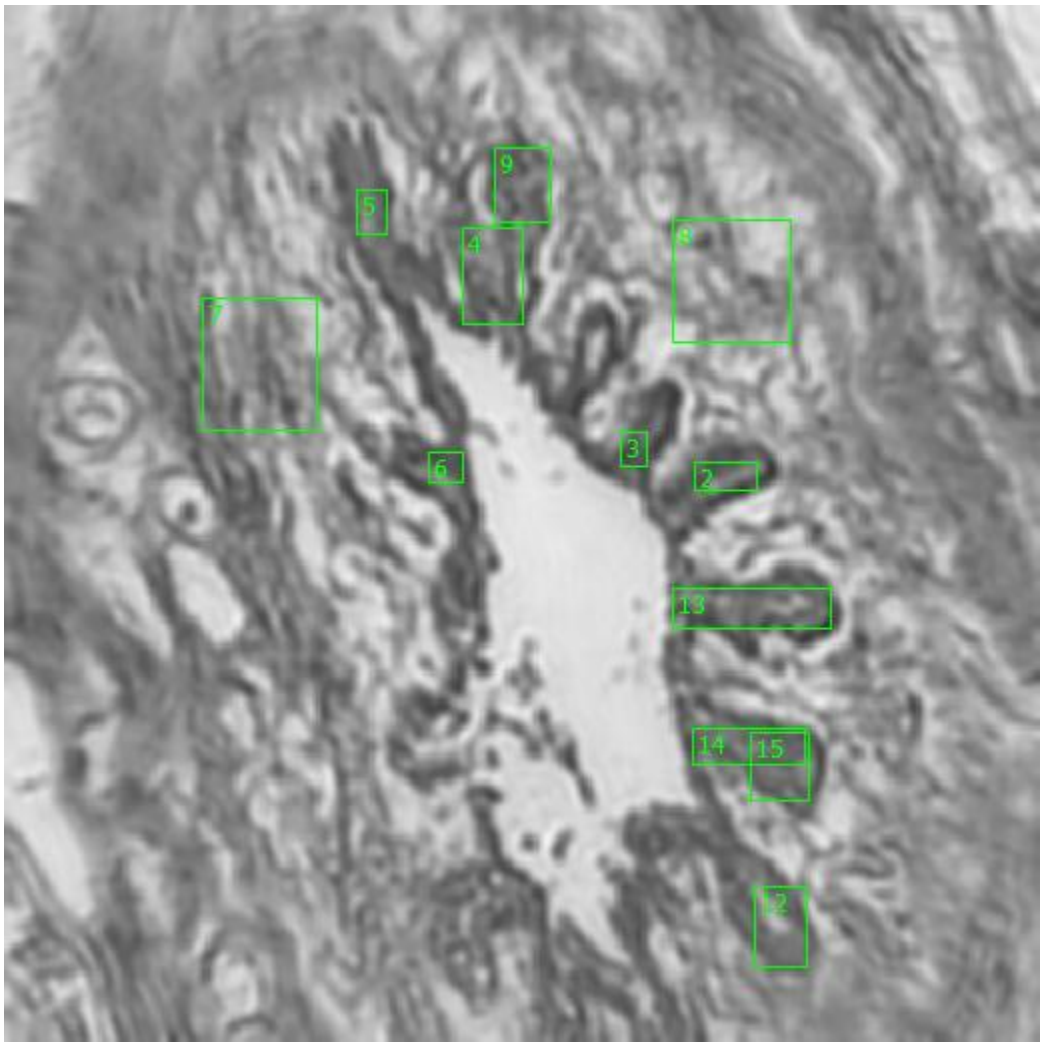



**Figure 0.111. Point T Regions of Interest**

**Regions 2, 12, 15:**
**Regions 3, 4, 5, 6, 9:**
**Regions 7, 8: breast**
**Regions 13, 14:**

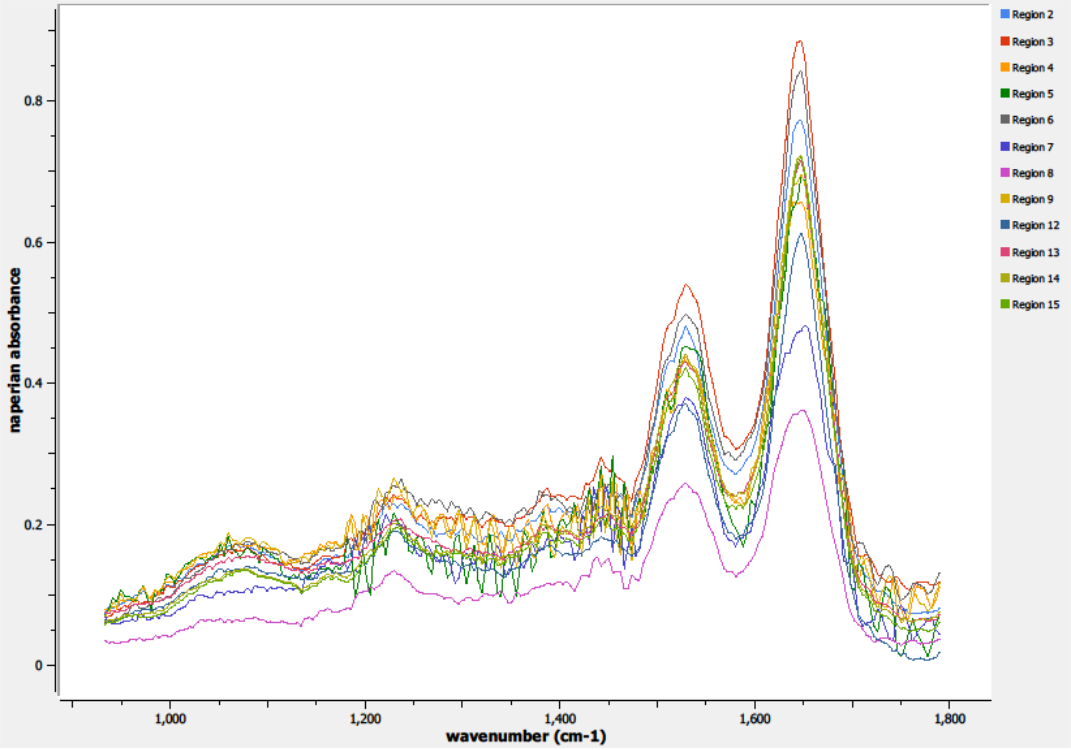



**Figure 0.112. Point T Spectra**

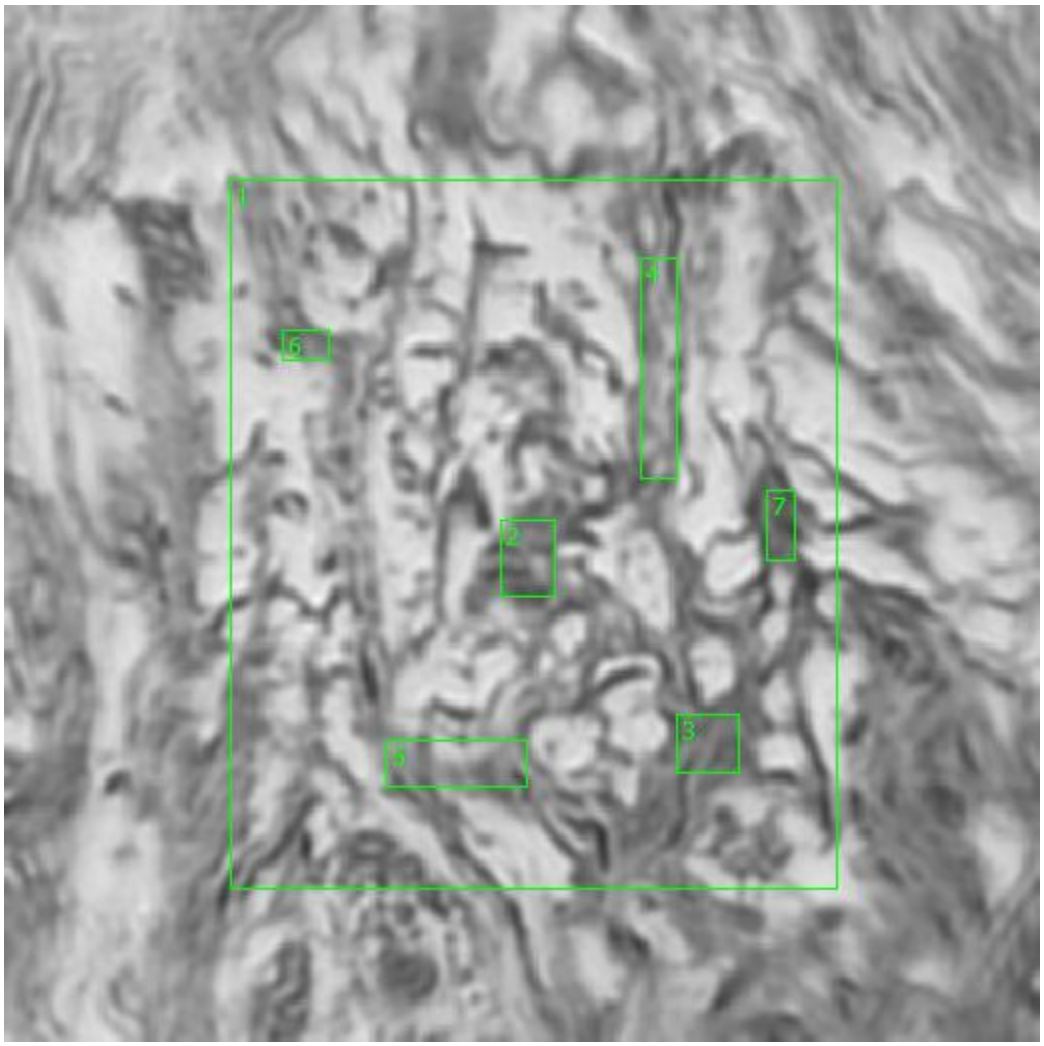



**Figure 0.113. Point V Regions of Interest**
**Region 1: breast**
**Regions 2, 3, 4, 5, 6, 7: concentrated region, breast**

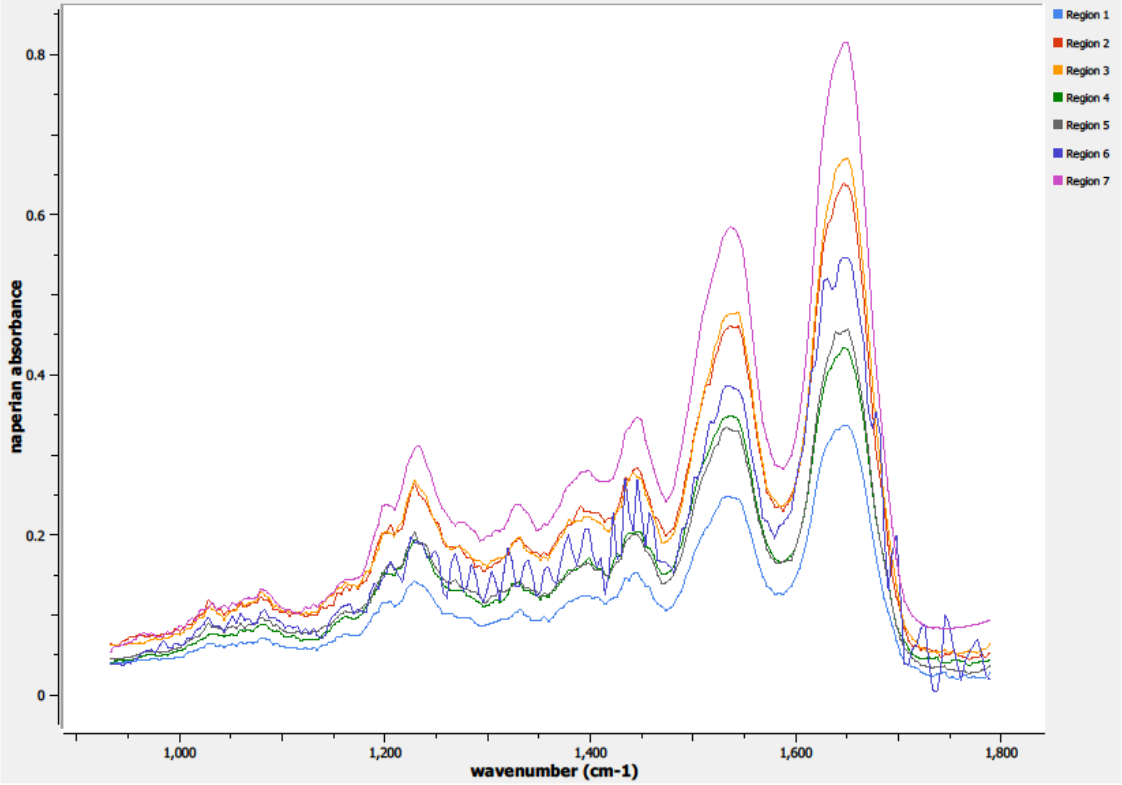



**Figure 0.114. Point V Spectra**

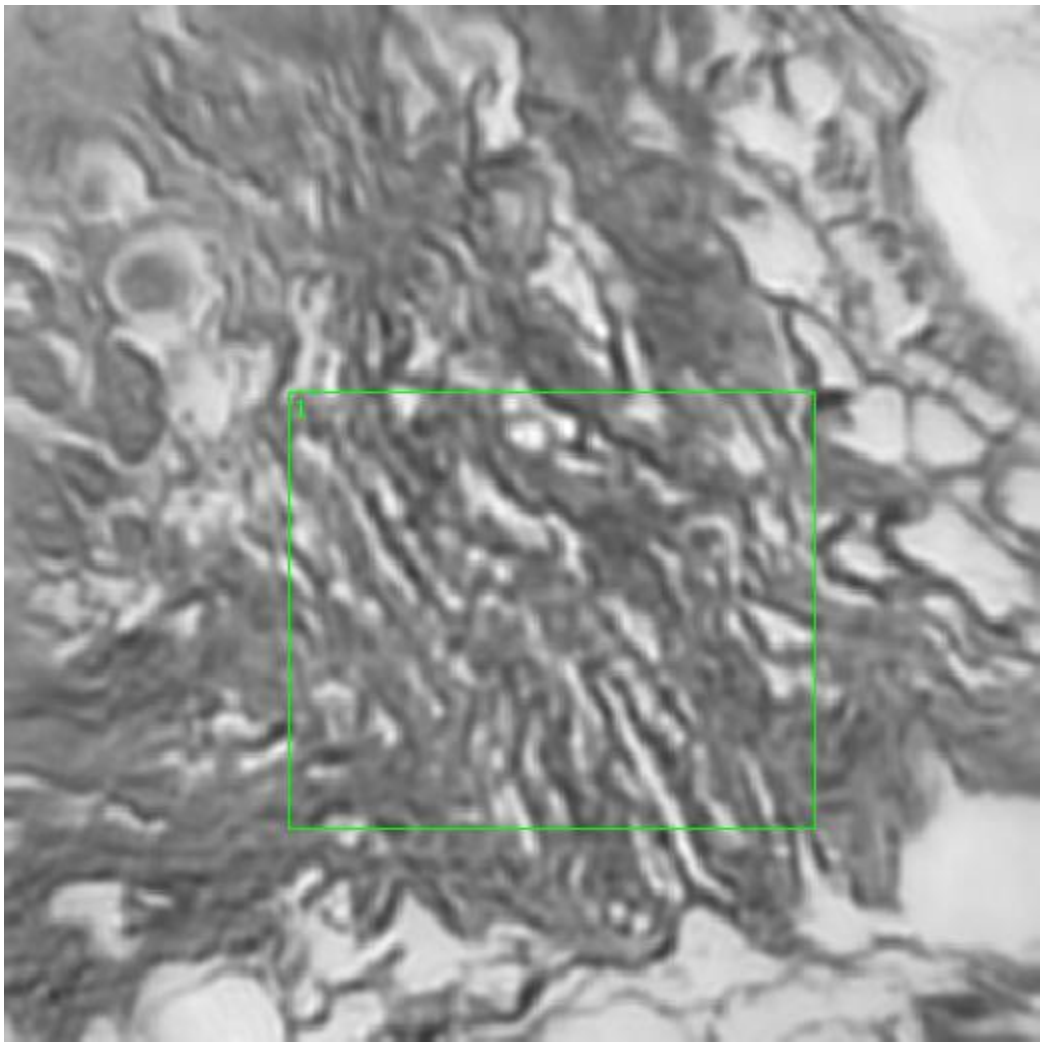


**Figure 0.115. Point W Regions of Interest**
**Region 1: breast**

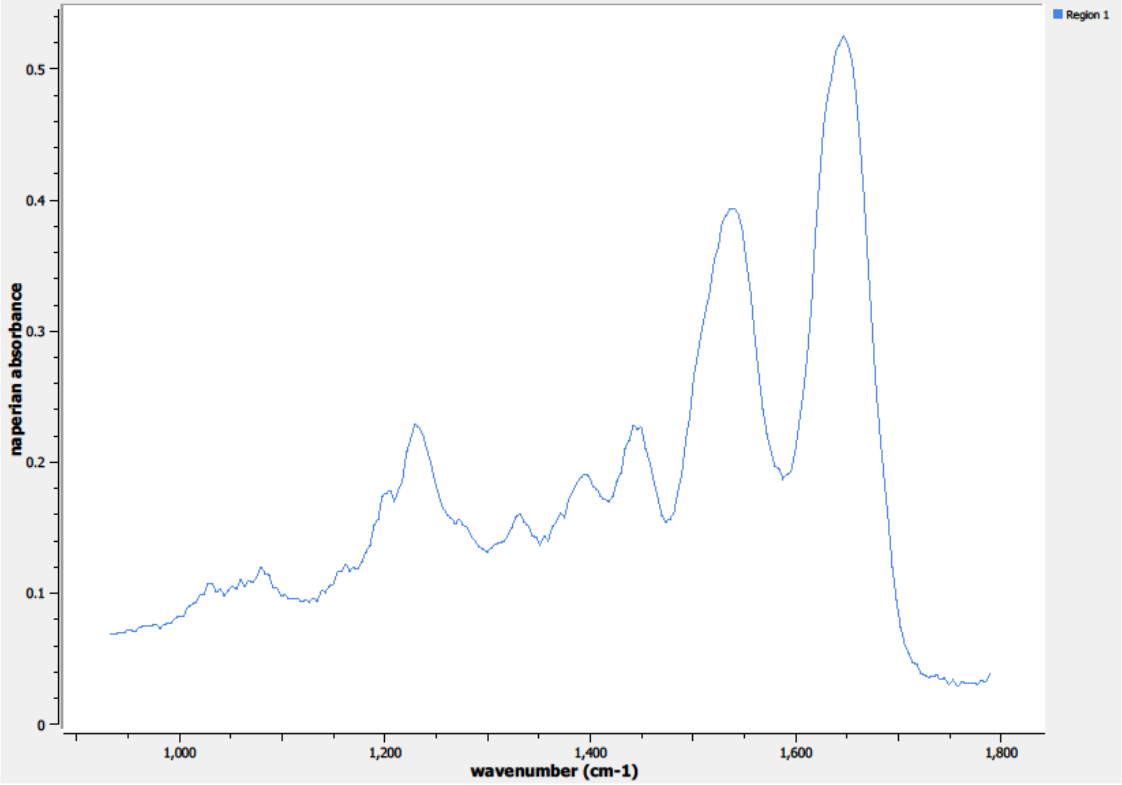



**Figure 0.116. Point W Spectra**

## A18. 13-06-A025a

| | |
|---|---|
| Unique ID | 13-06-A025a |
| Age | 59 |
| Race | white |
| Location of Sample Collection | left breast |
| Notes of Interest | not normal tissue sample, tissue sample too dense, unused in data results |

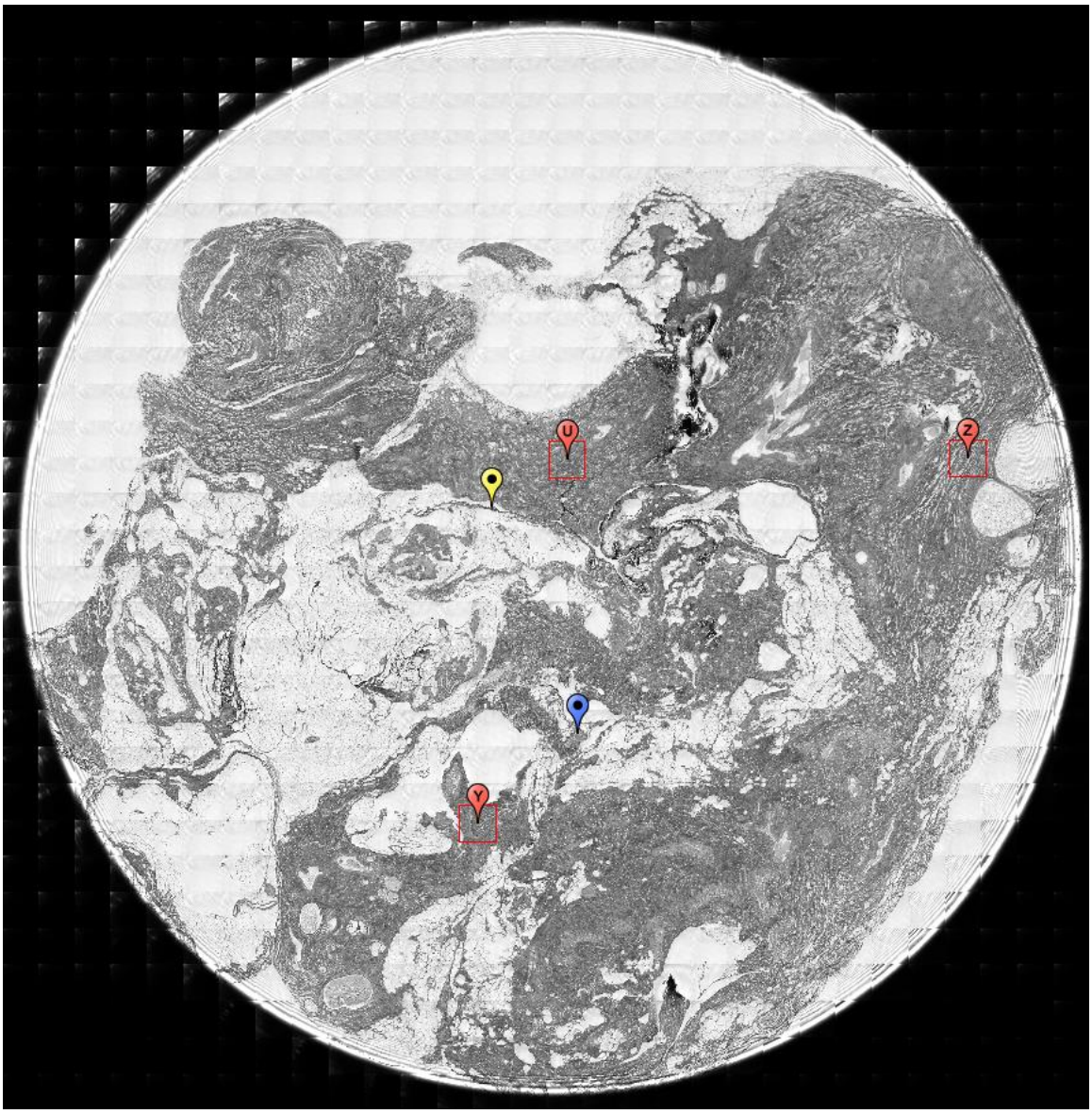



**Figure 0.117. Specimen 13-06-A025a, normal breast tissue.**

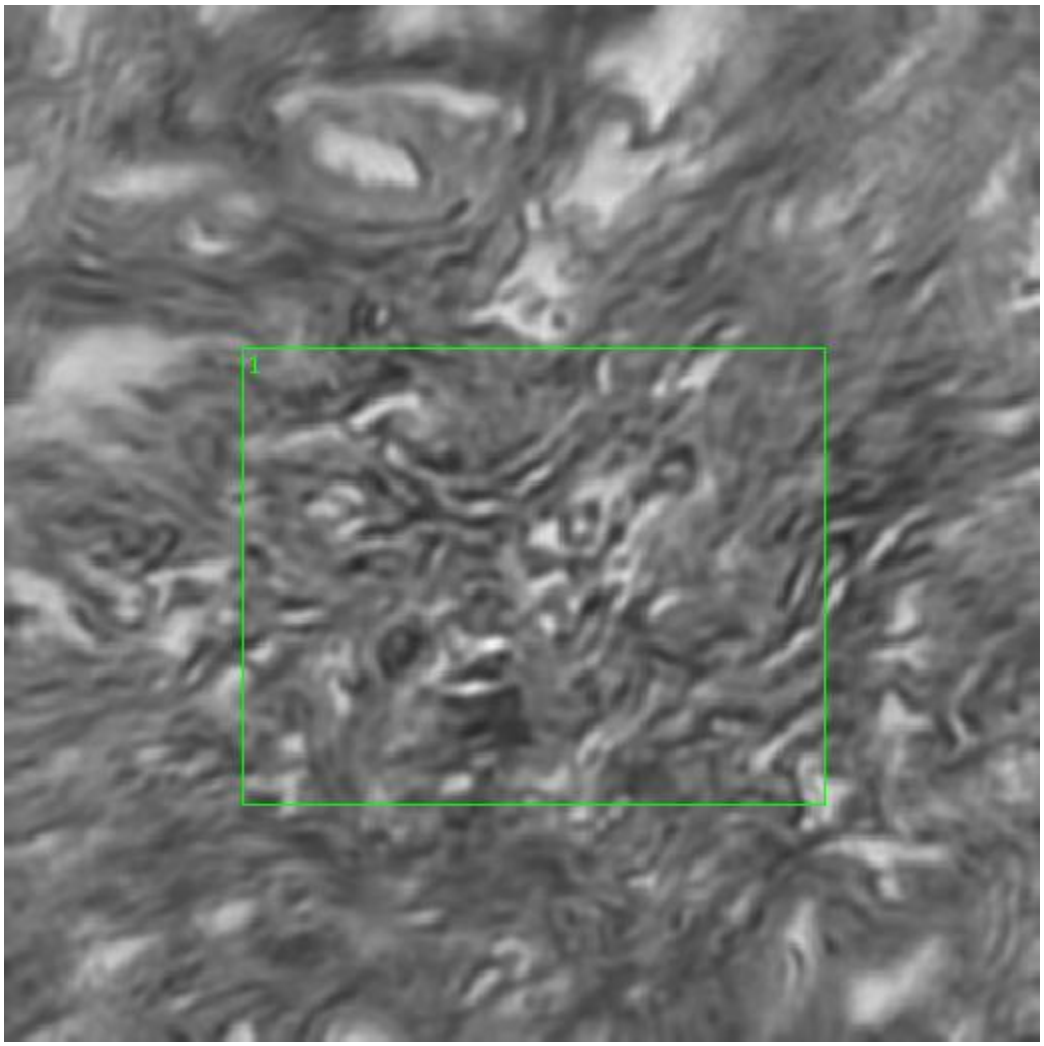



**Figure 0.118. Point U Regions of Interest**
**Region 1: left breast**

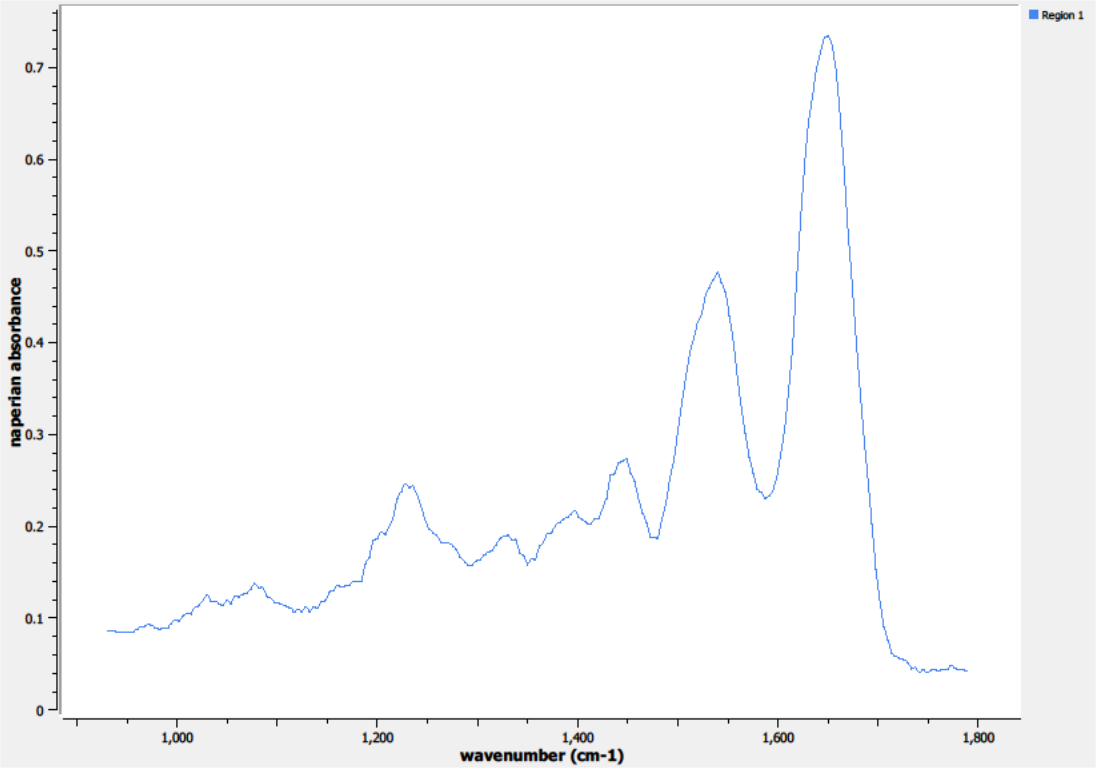



**Figure 0.119. Point U Spectra**

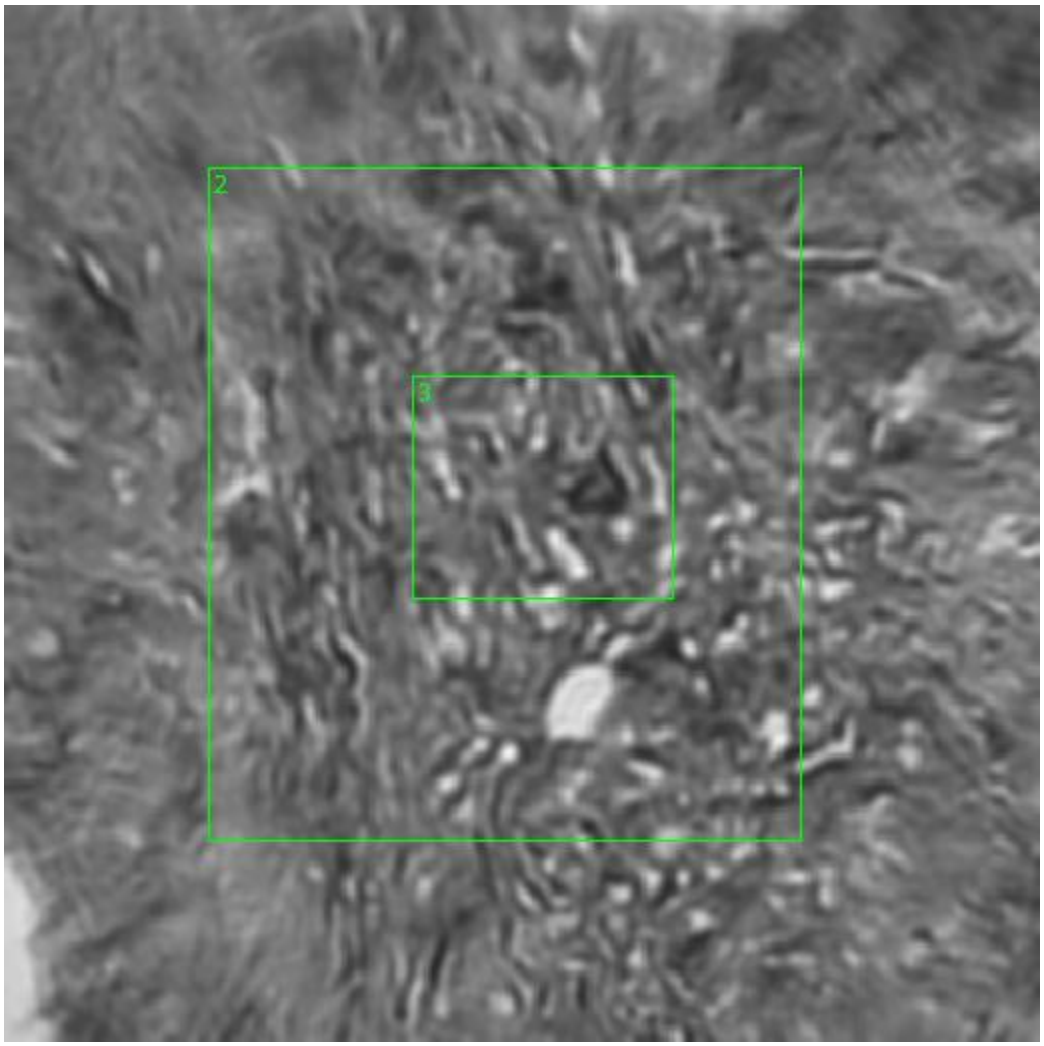



**Figure 0.120. Point Y Regions of Interest**
**Region 1: left breast**
**Region 2: concentrated region, left breast**

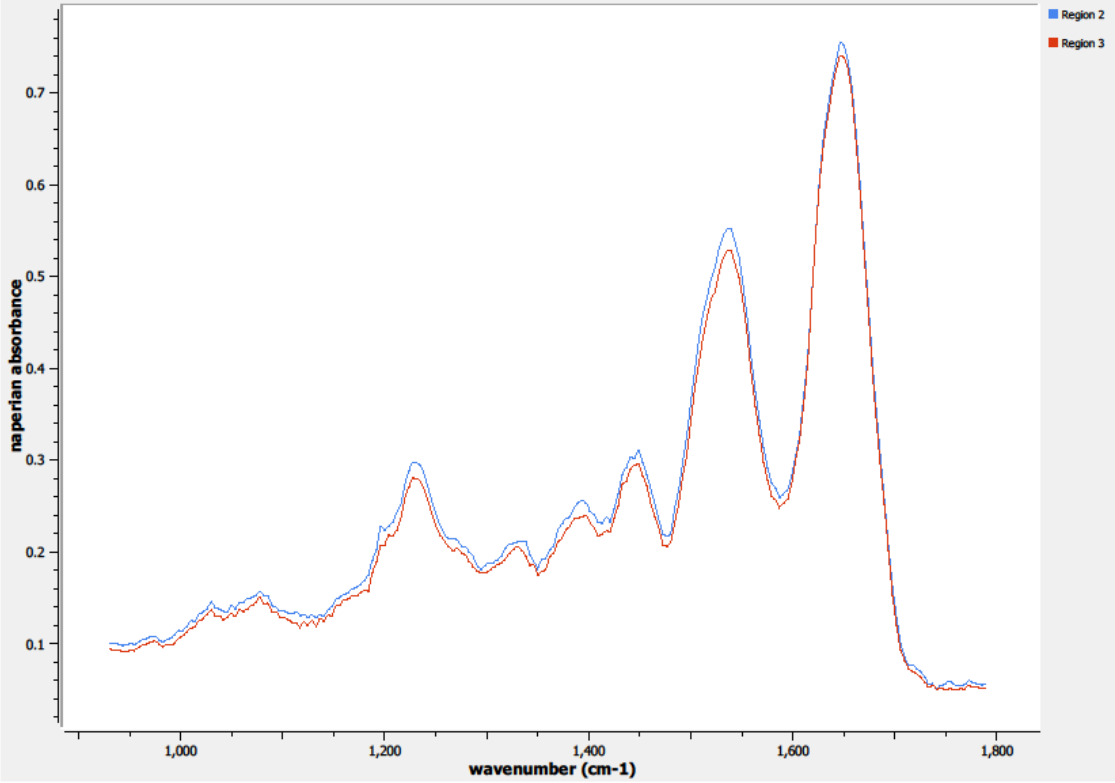



**Figure 0.121. Point Y Spectra**

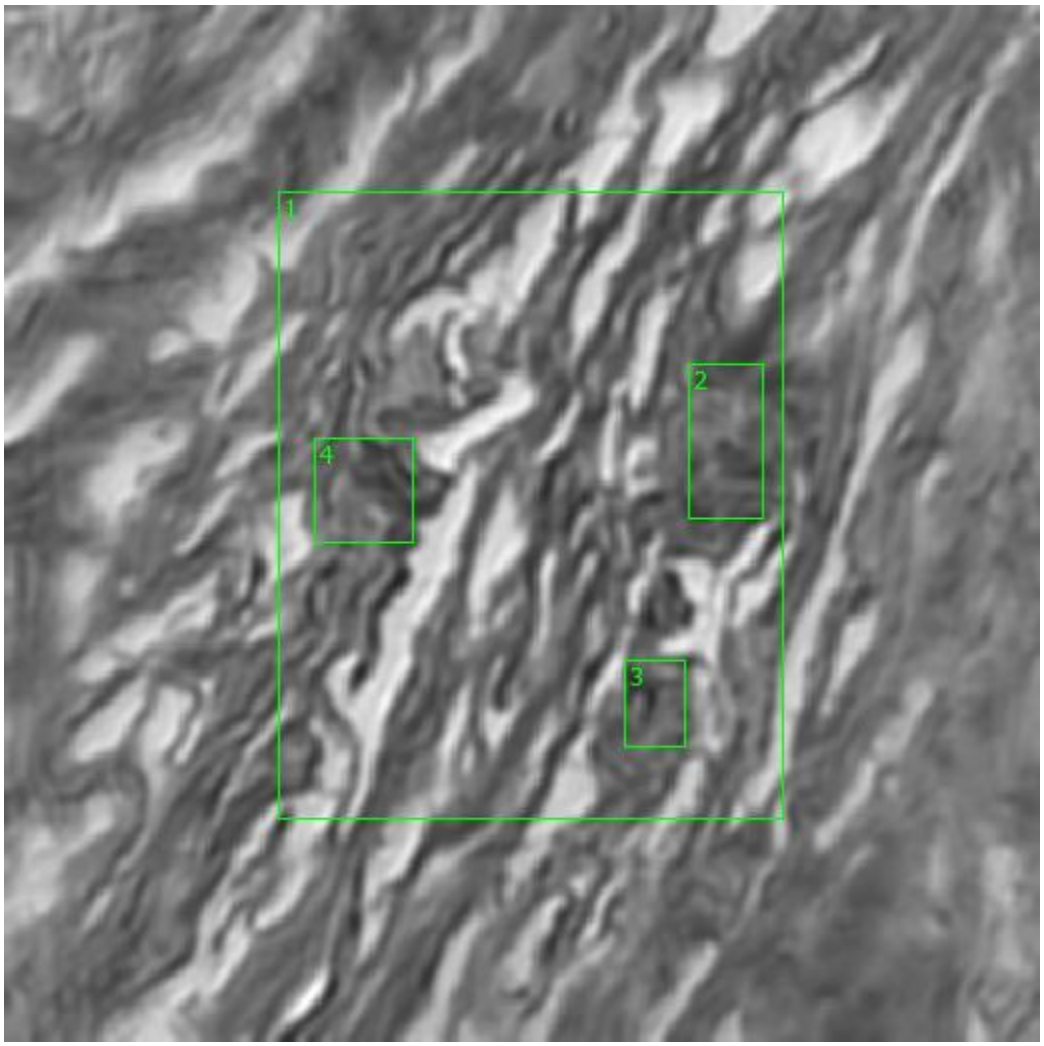



**Figure 0.122. Point Z Regions of Interest**
**Region 1: left breast**
**Regions 2, 3, 4: concentrated regions, left breast**

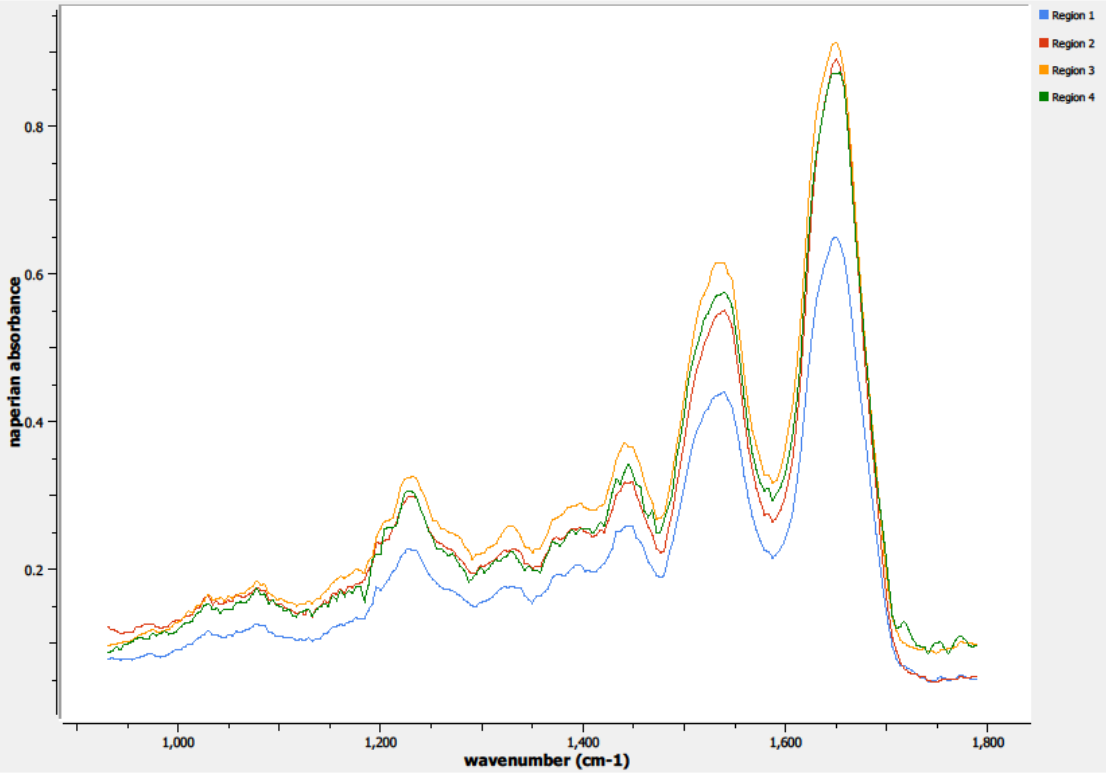



**Figure 0.123. Point Z Spectra**

# Appendix D

# Breast Cancer

**A1.08-01-A132u**

| | |
|---|---|
| Unique ID | 08-01-A132u |
| Age | 55 |
| Race | white |
| Location of Sample Collection | breast, phylloides |
| Notes of Interest | malignant cystosarcoma phylloides, %T (100) / %N (50) |

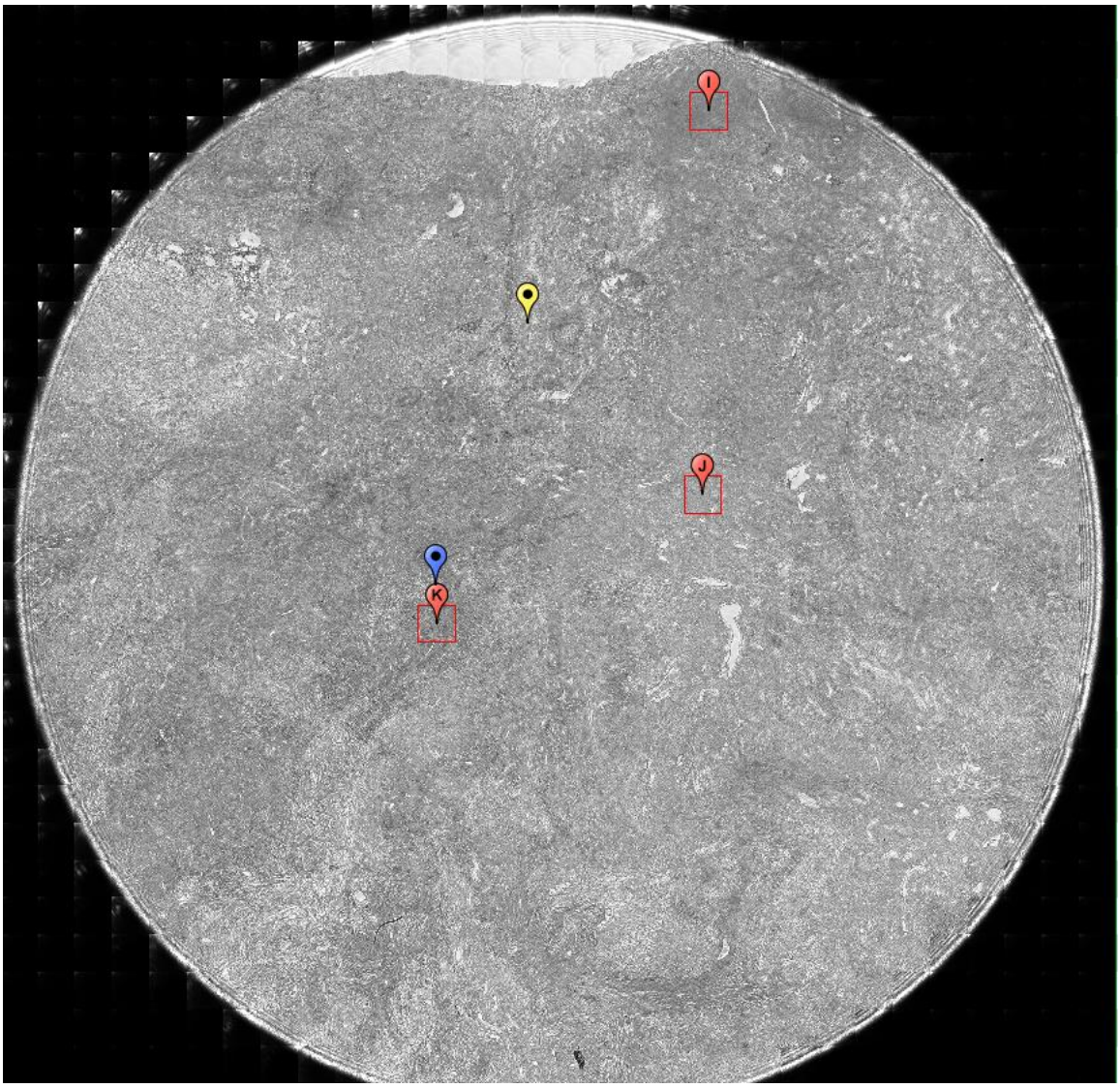



**Figure 0.1. Specimen 08-01-A132u, breast cancer tissue**

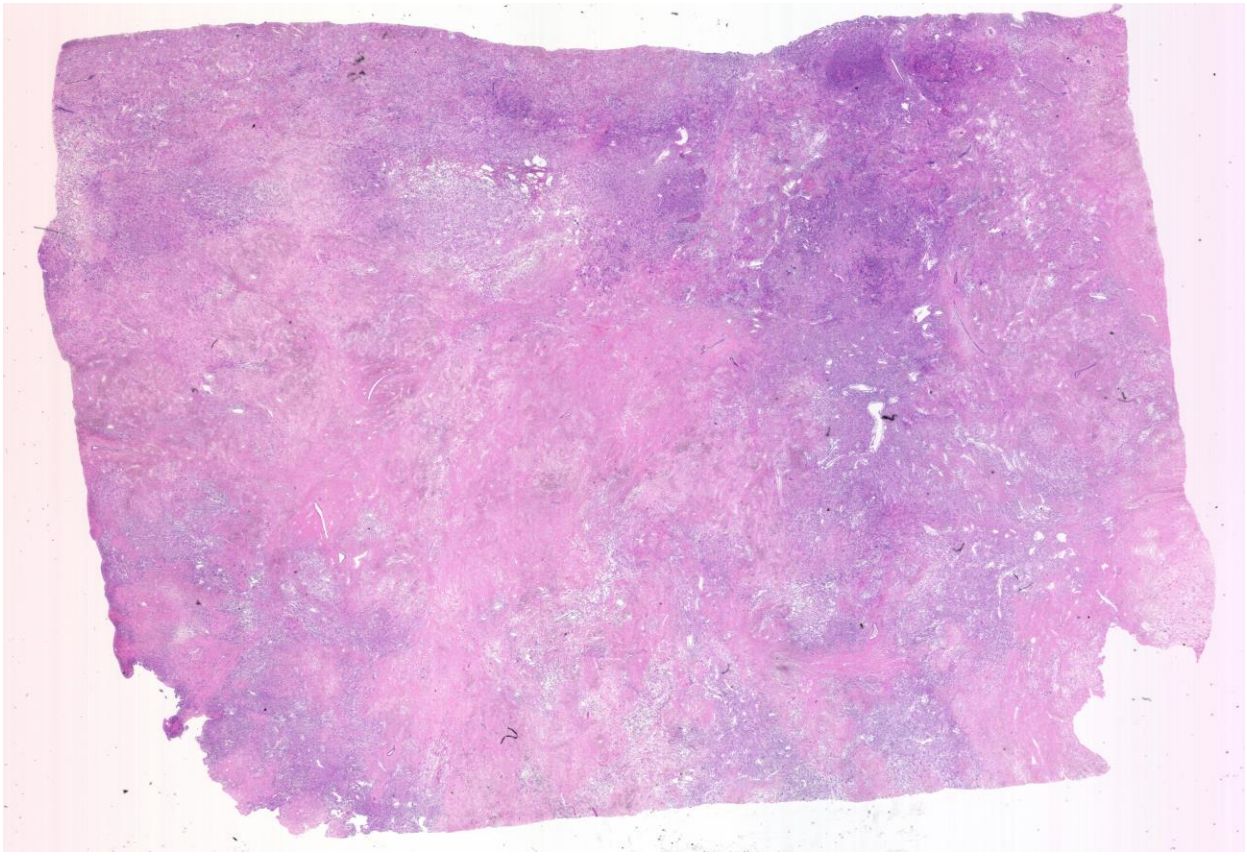


**Figure 0.2. Specimen 08-01-A132u, H&E stain, malignant cystosarcoma phylloides of the breast**

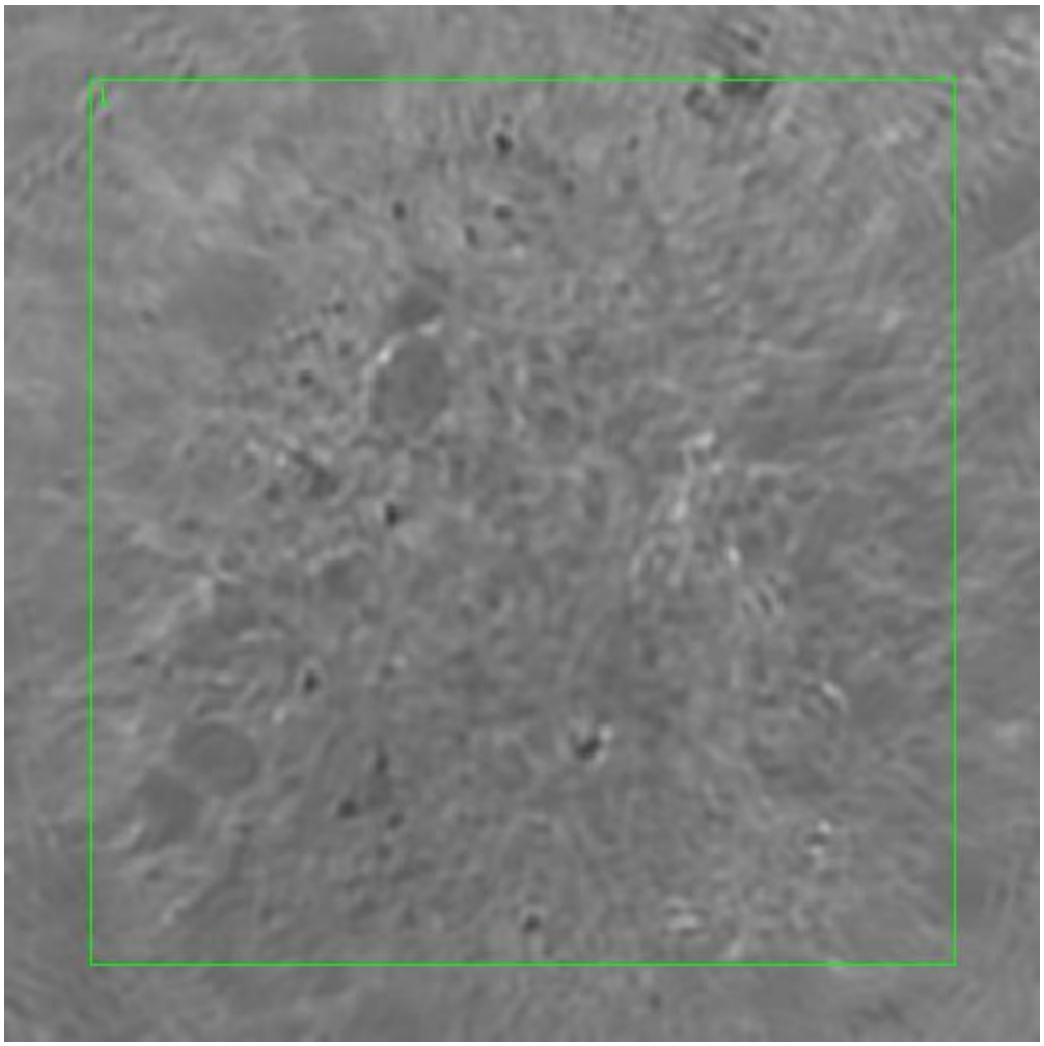


**Figure 0.3. Point I Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

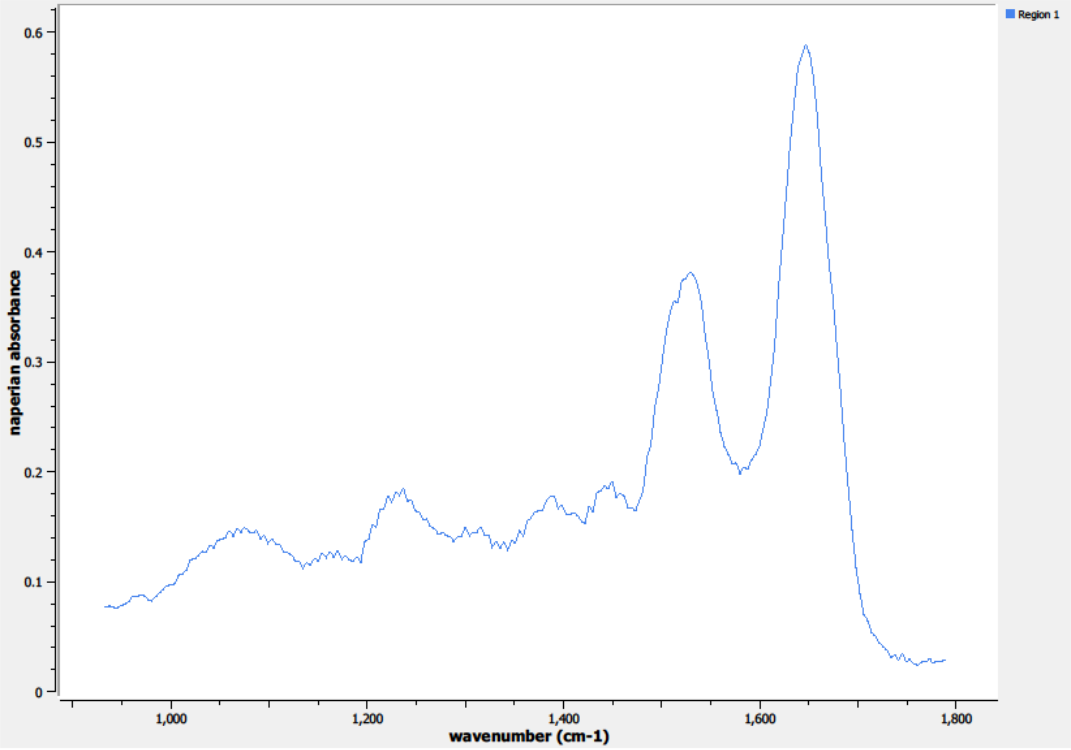



**Figure 0.4. Point I Spectra**

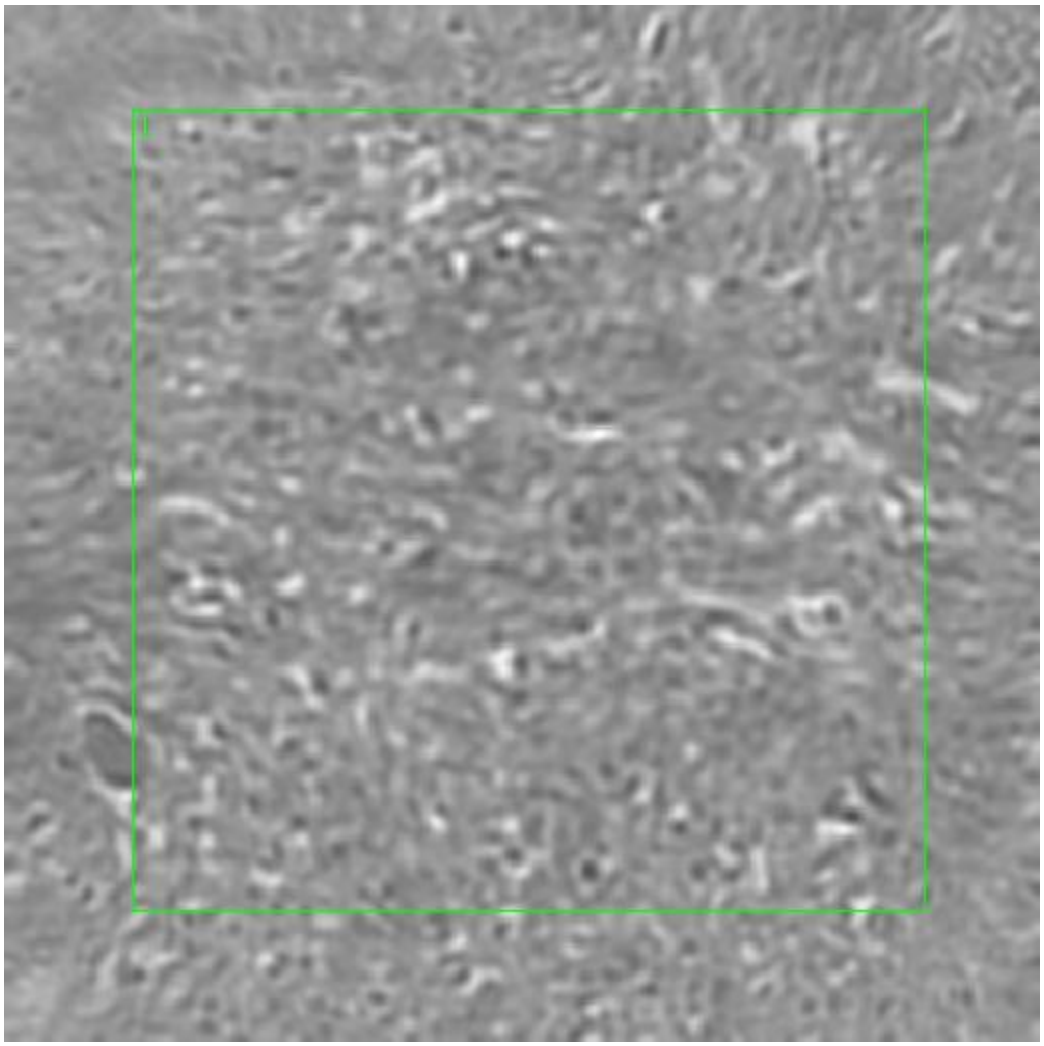


**Figure 0.5. Point J Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

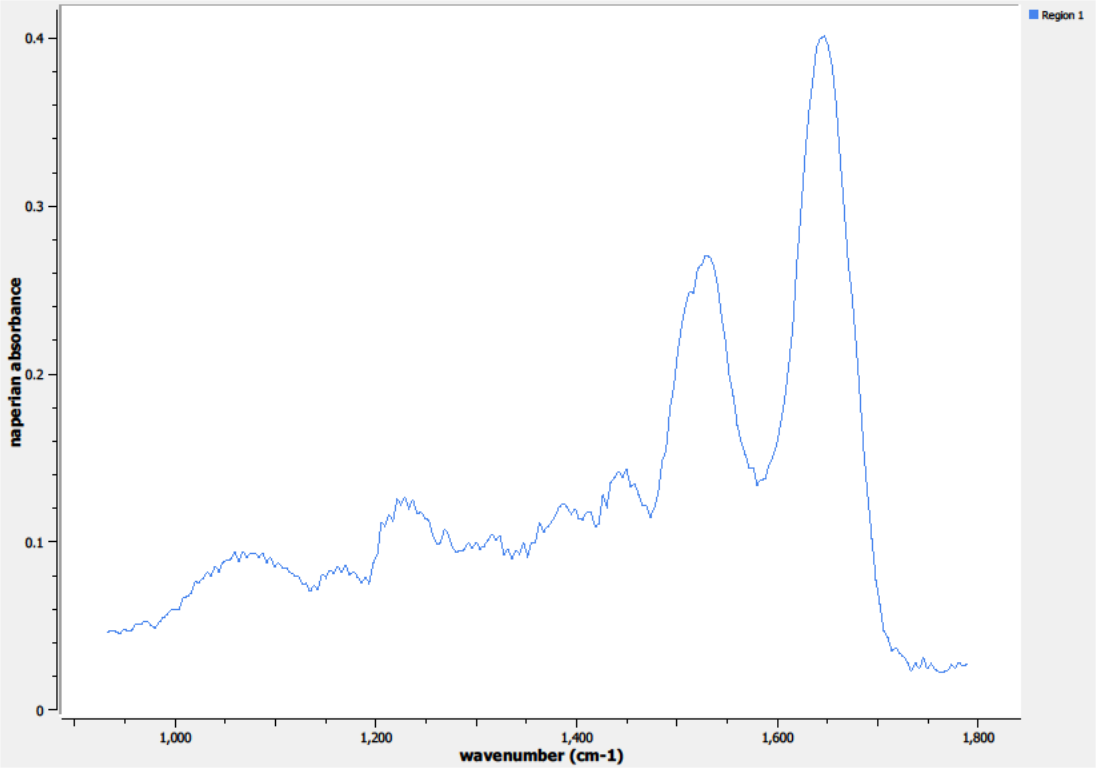



**Figure 0.6. Point J Spectra**

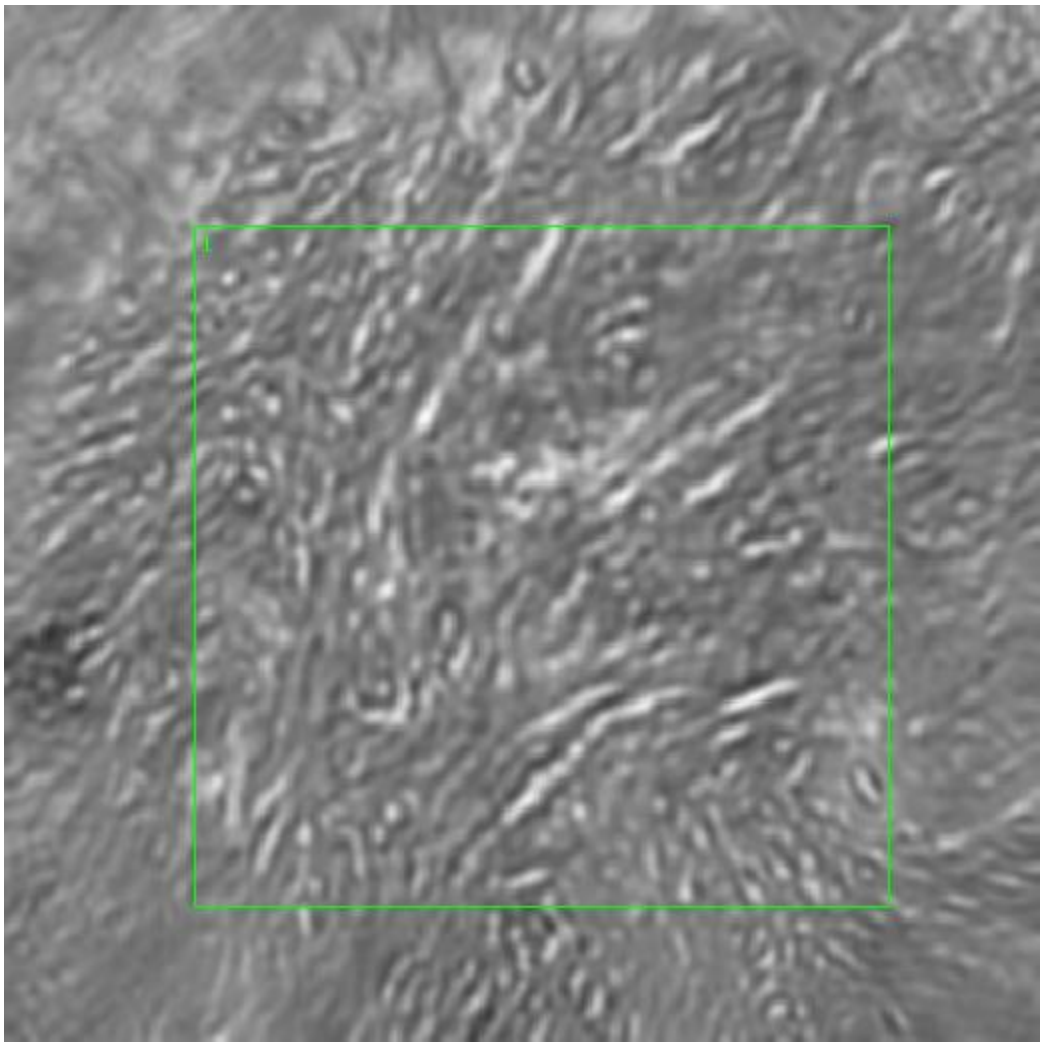


**Figure 0.7. Point K Regions of Interest**
**Region 1: breast, malignant cystosarcoma phylloides**

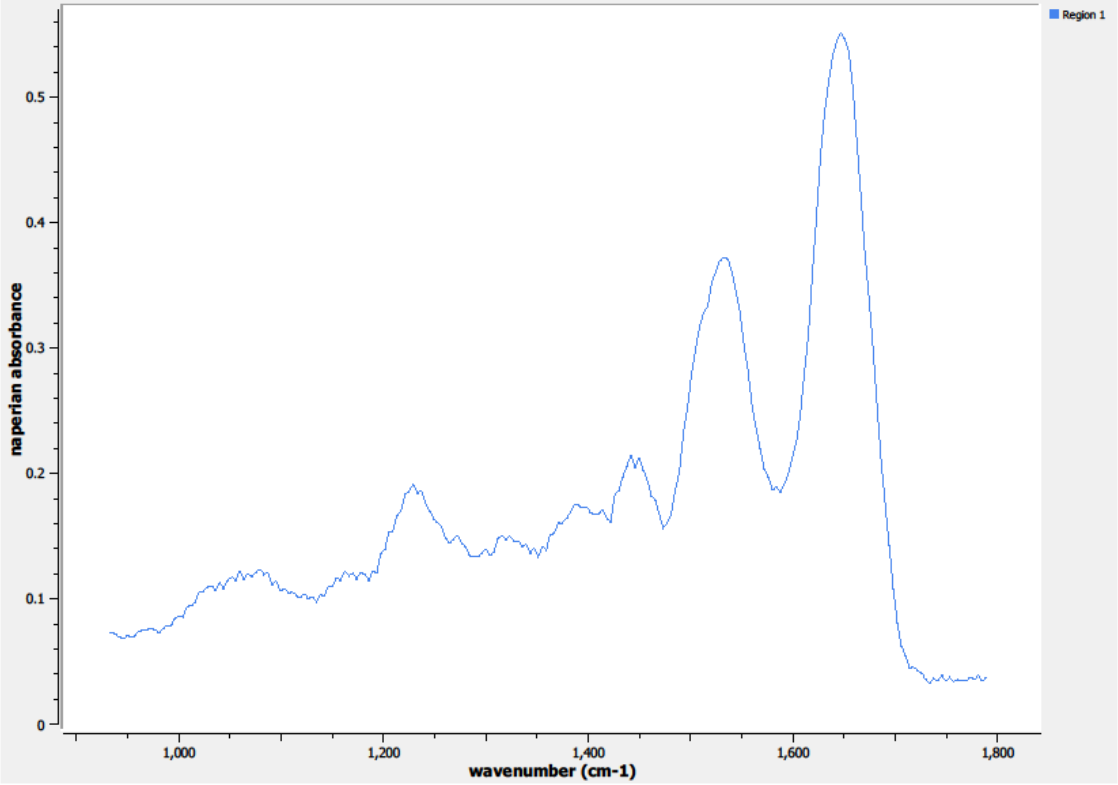



**Figure 0.8. Point K Spectra**

## A2.08-12-A025o

| | |
|---|---|
| Unique ID | 08-12-A025o |
| Age | 45 |
| Race | Asian |
| Location of Sample Collection | breast, ductal |
| Notes of Interest | malignant ductal adenocarcinoma, inflammatory carcinoma, %T (80) / %N (60) |

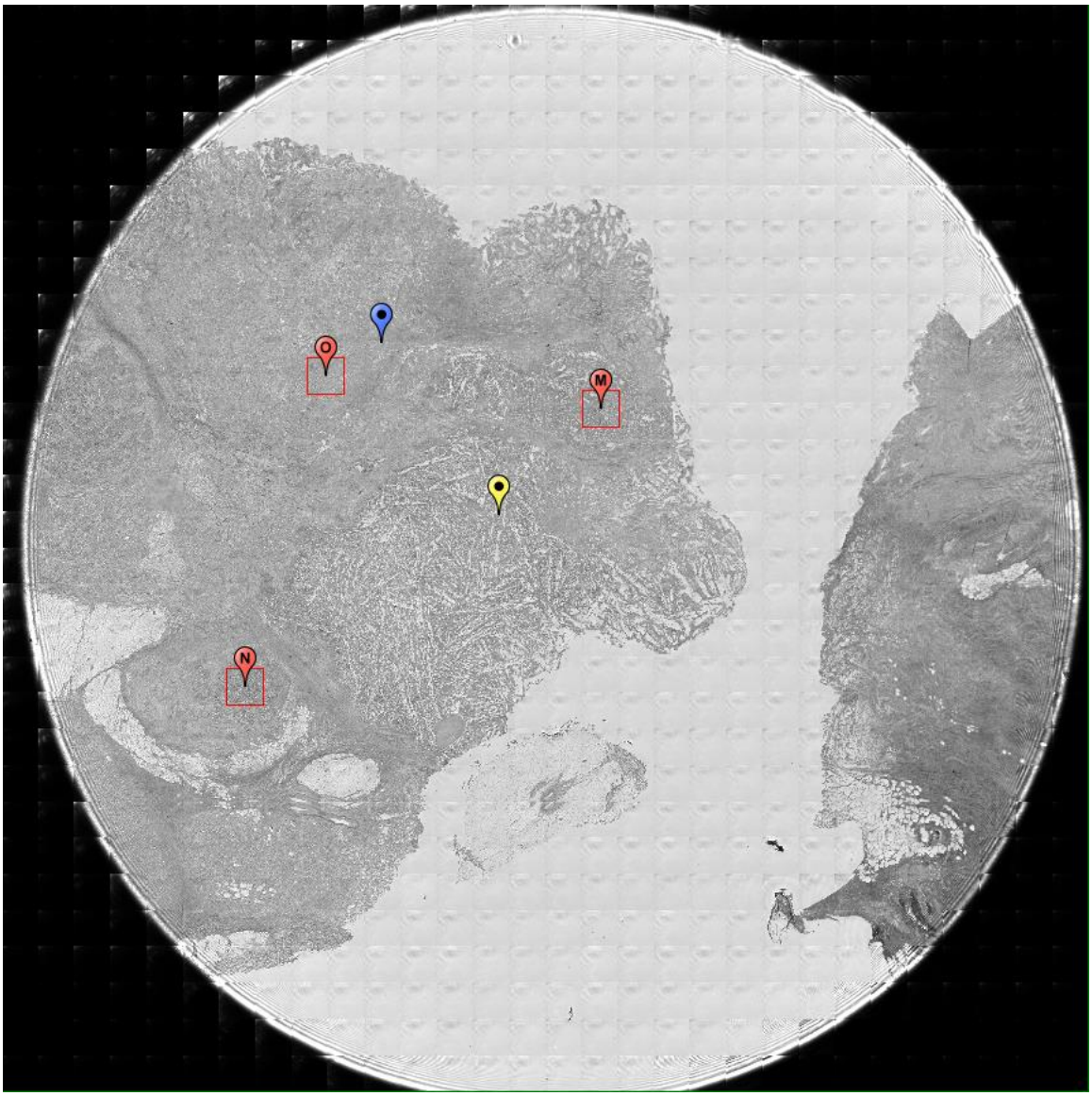



**Figure 0.9. Specimen 08-12-A025o, breast cancer tissue**

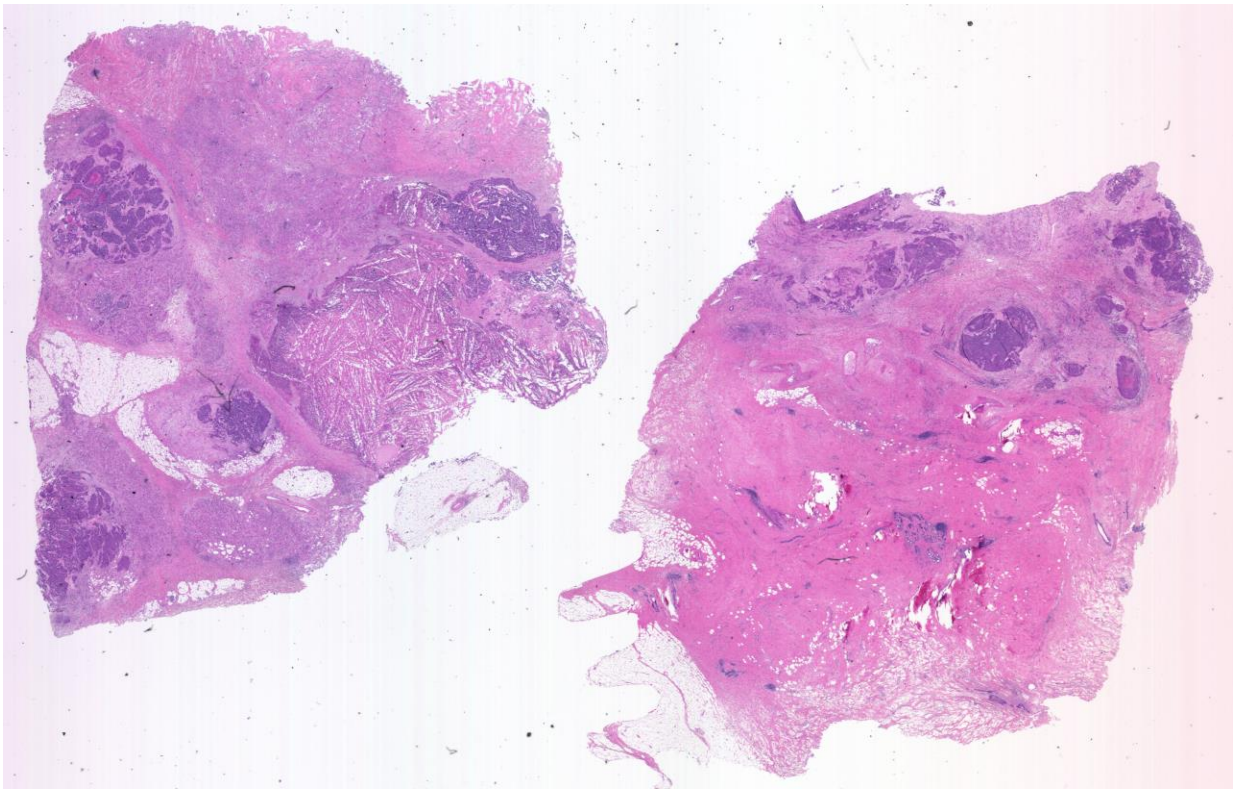


**Figure 0.10. Specimen 08-12-A025o, H&E stain, malignant ductal adenocarcinoma of the breast**

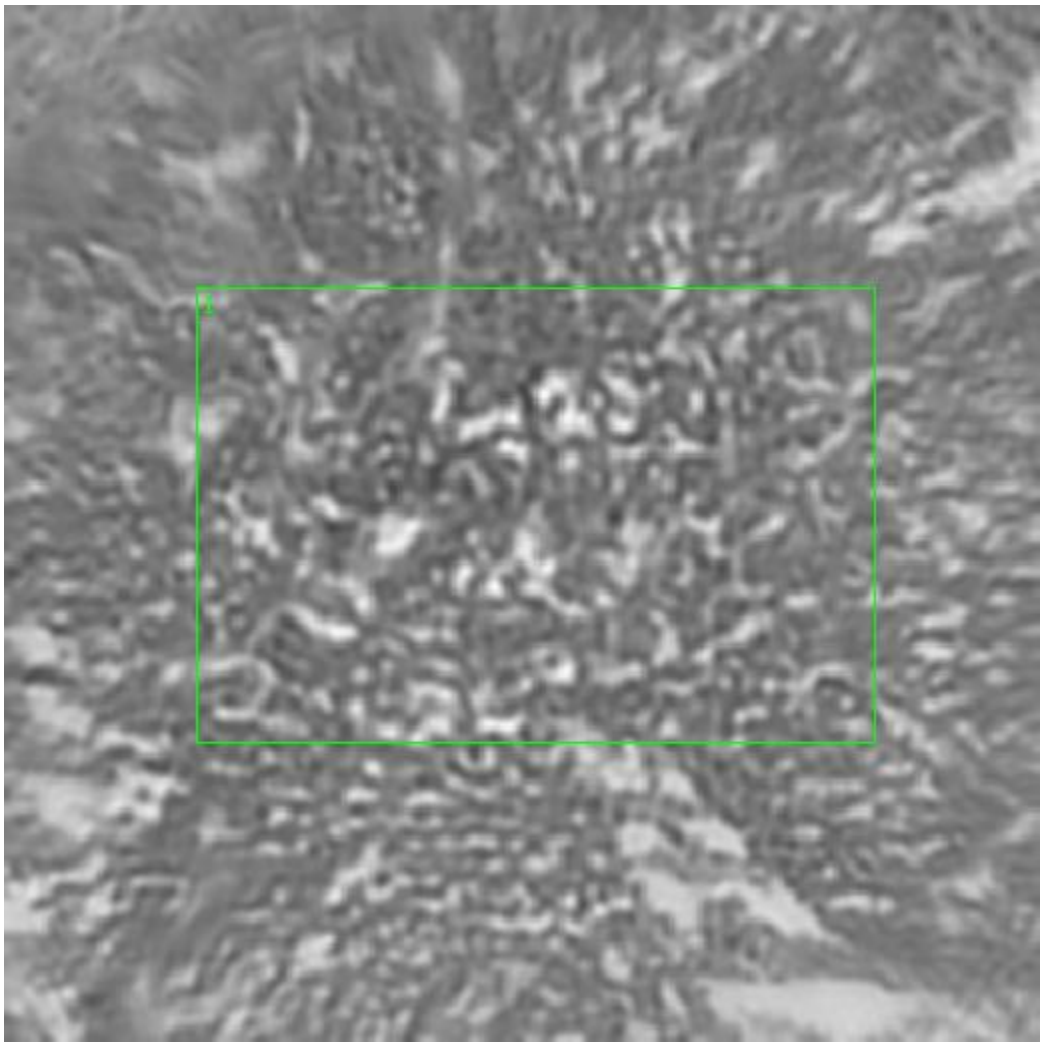


**Figure 0.11. Point M Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

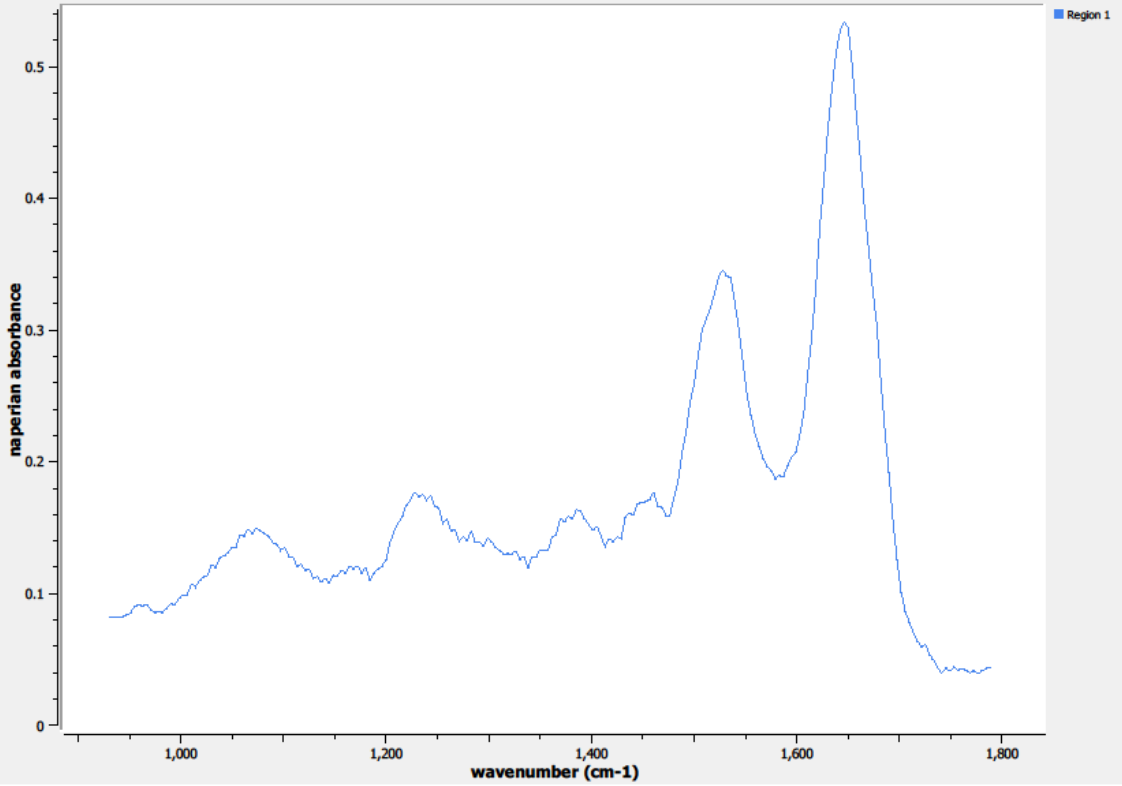



**Figure 0.12. Point M Spectra**

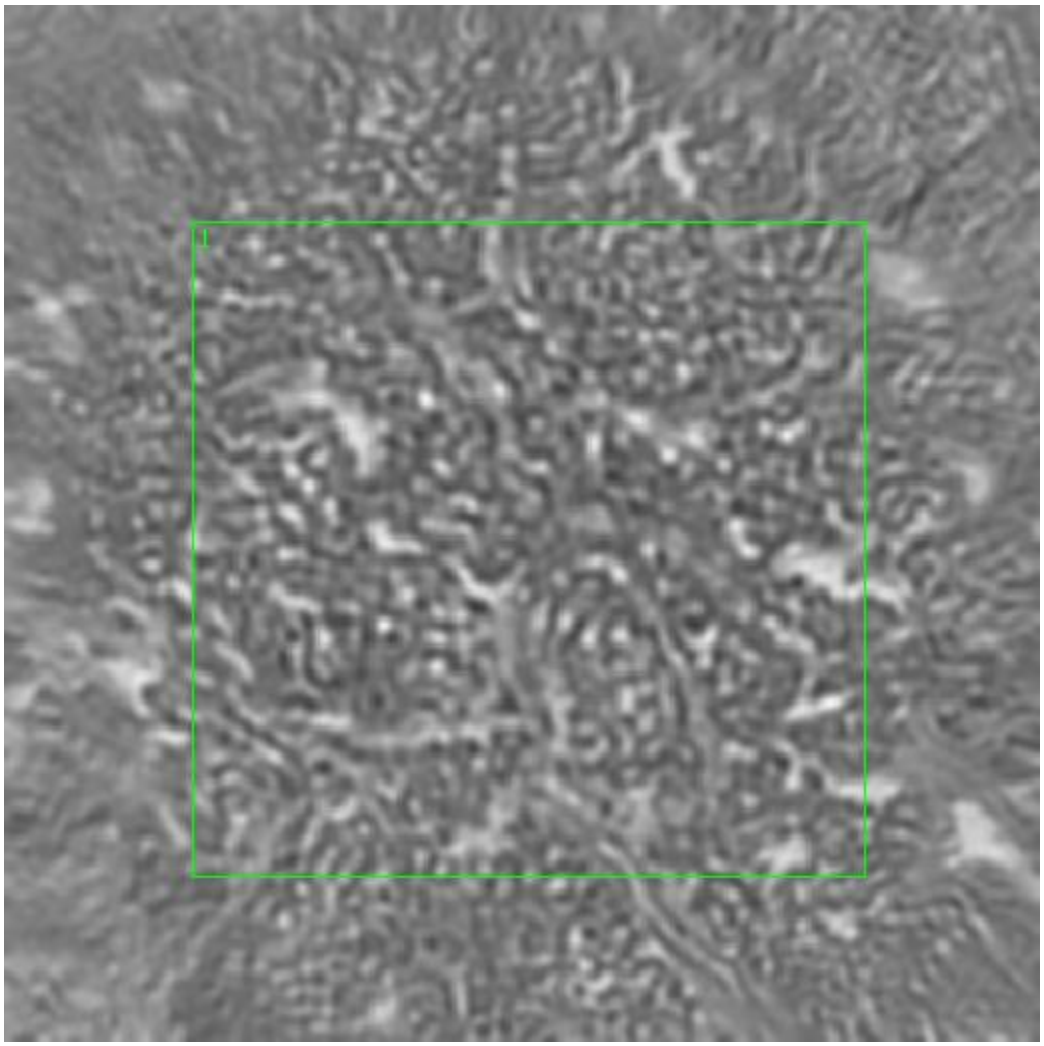


**Figure 0.13. Point N Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

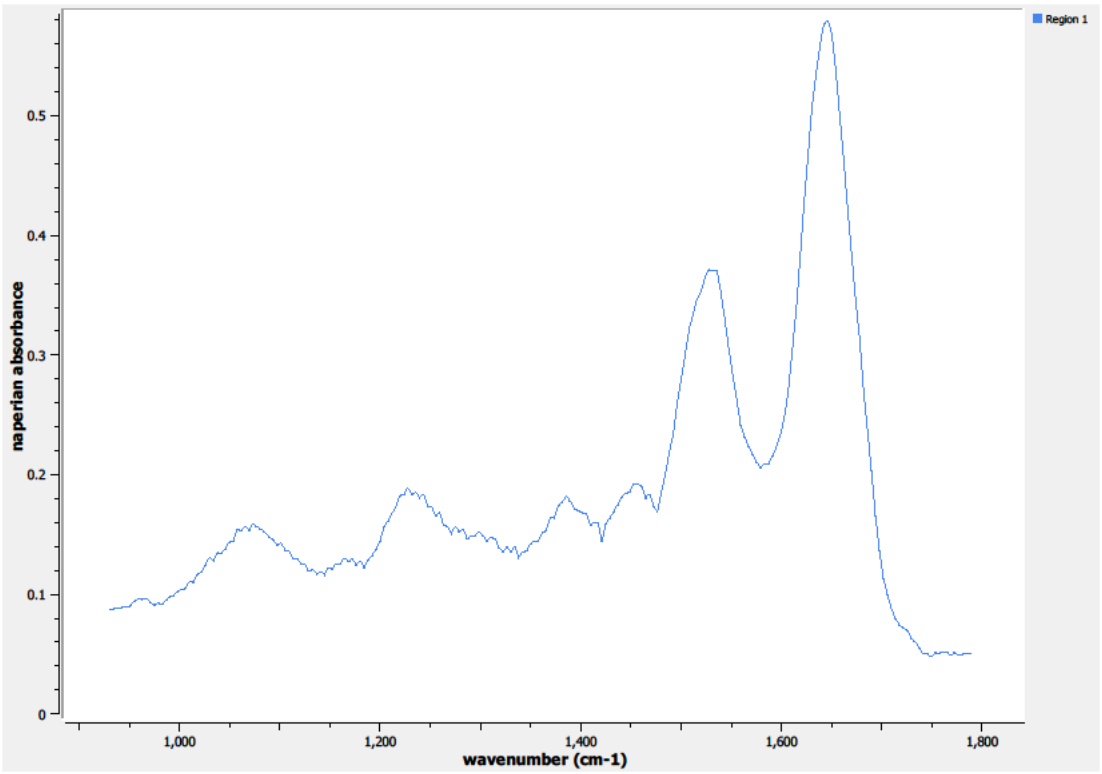



**Figure 0.14. Point N Spectra**

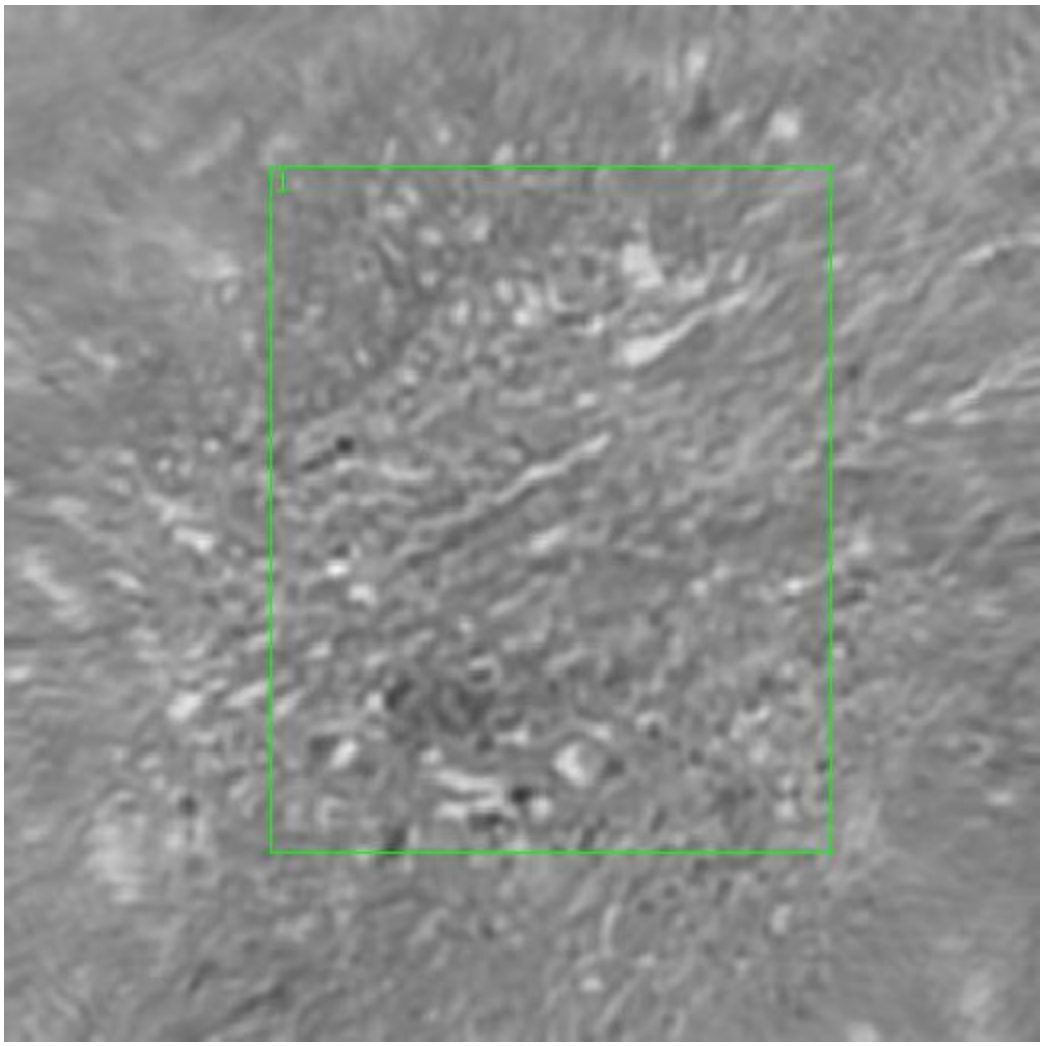


**Figure 0.15. Point O Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

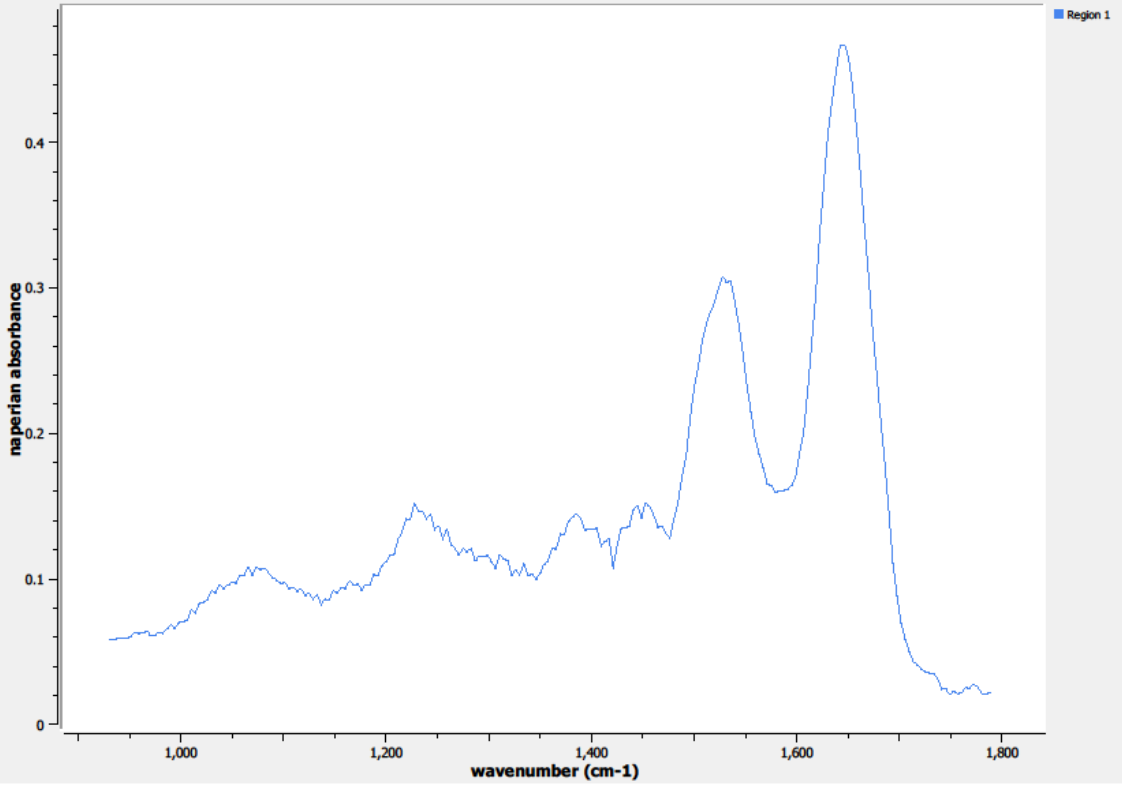



**Figure 0.16. Point O Spectra**

## A3.10-10-A177m

| | |
|---|---|
| Unique ID | 10-10-A177m |
| Age | 60 |
| Race | white |
| Location of Sample Collection | breast, ductal |
| Notes of Interest | malignant ductal adenocarcinoma, all receptors negative (triple negative), %T (100) / %N (80) |

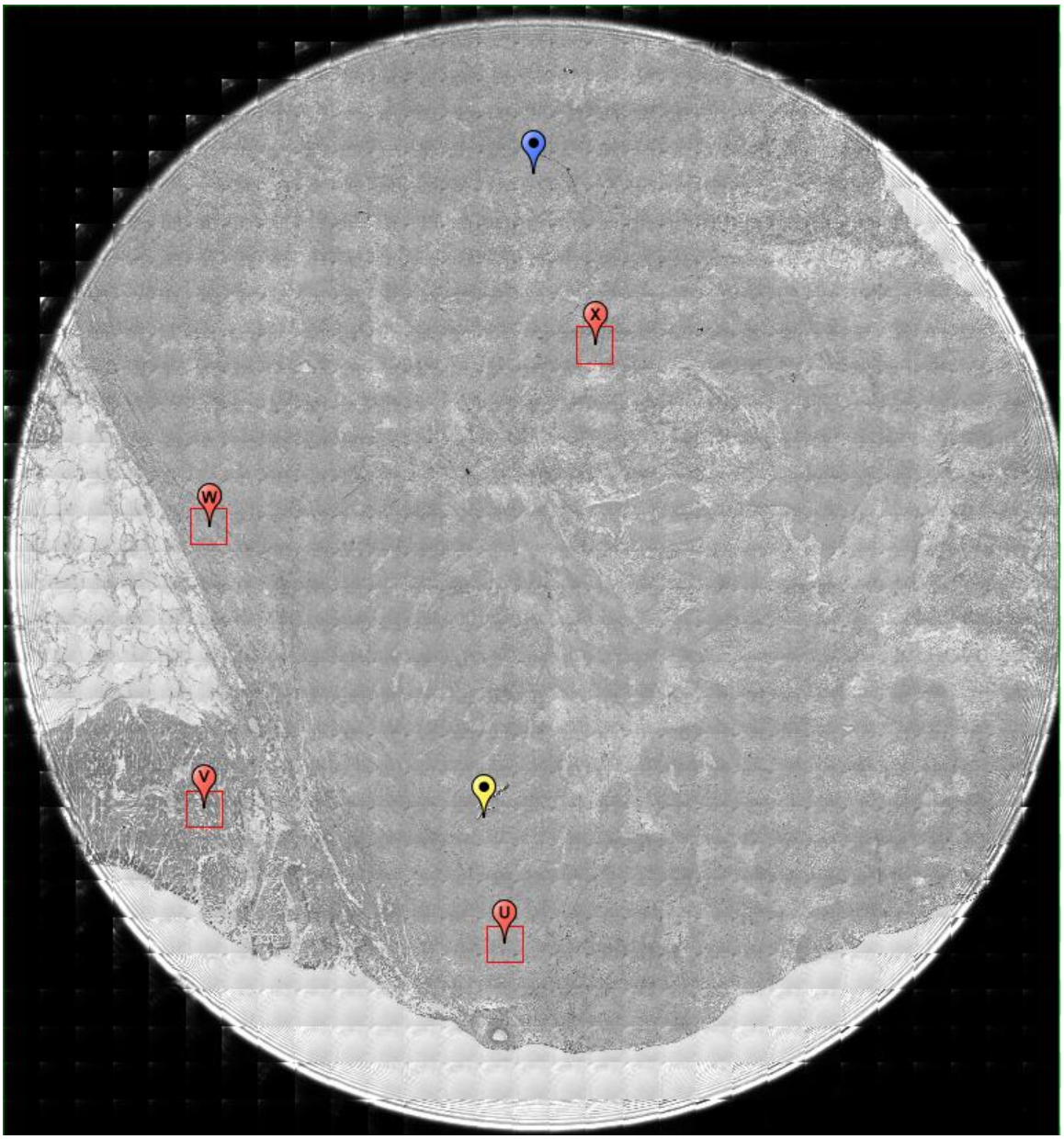



**Figure 0.17. Specimen 10-10-A177m, breast cancer tissue**

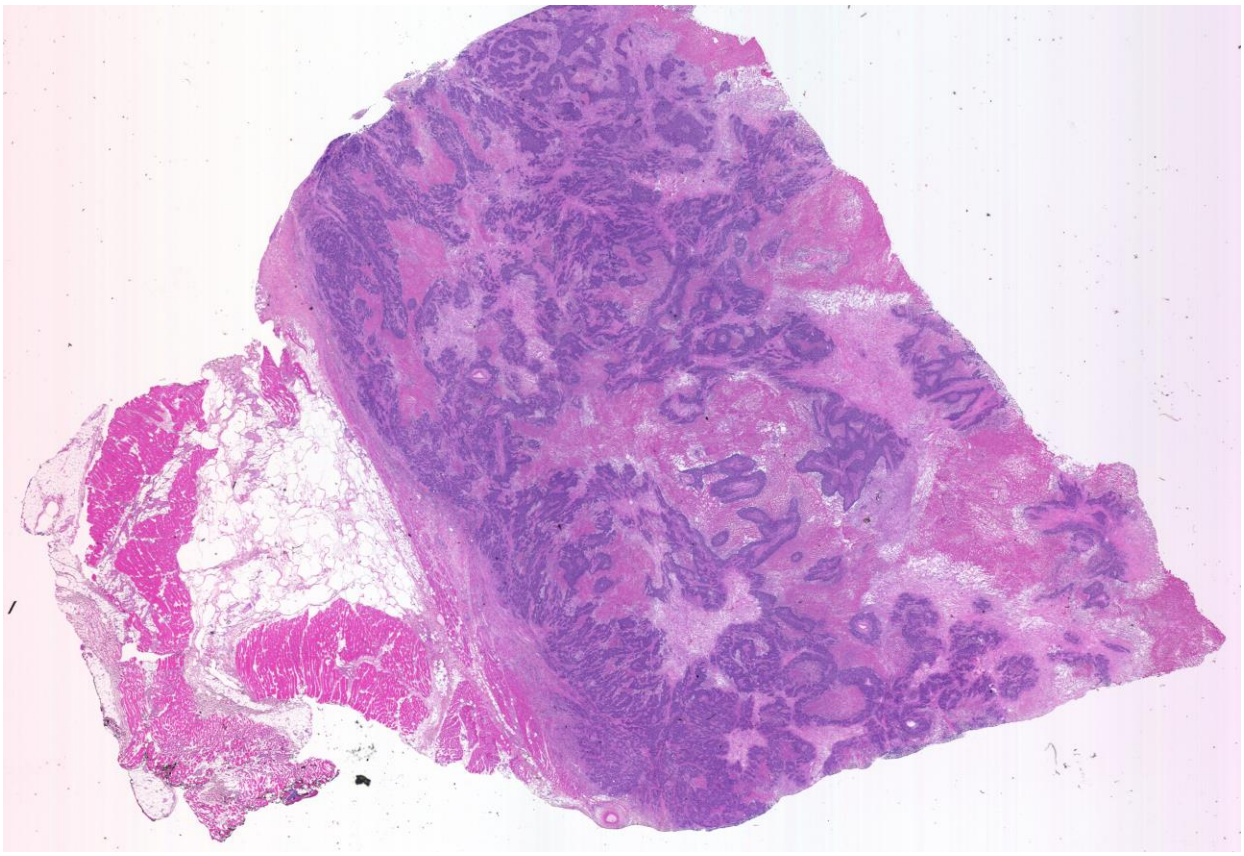


**Figure 0.18. Specimen 10-10-A177m, H&E stain, malignant ductal adenocarcinoma of the breast, triple negative**

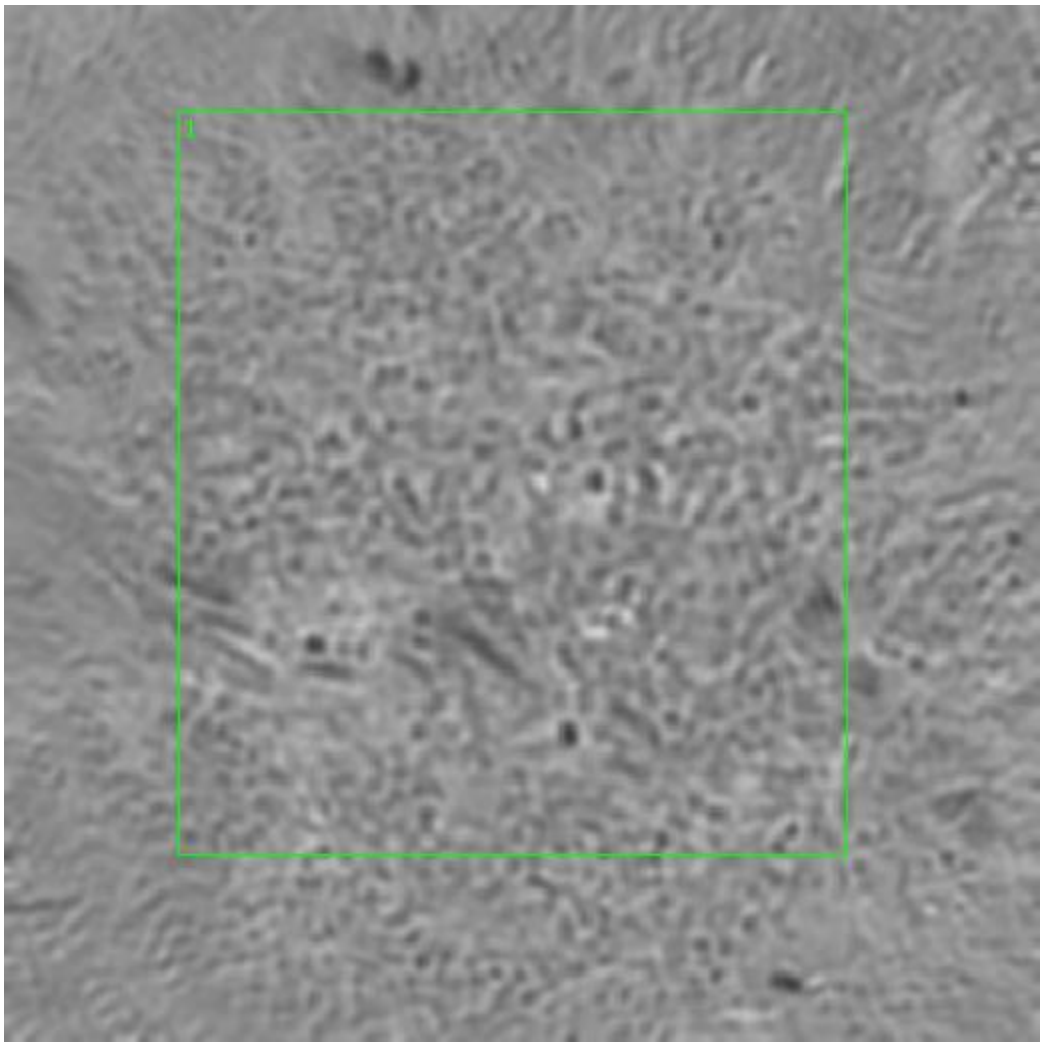



**Figure 0.19. Point U Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative**

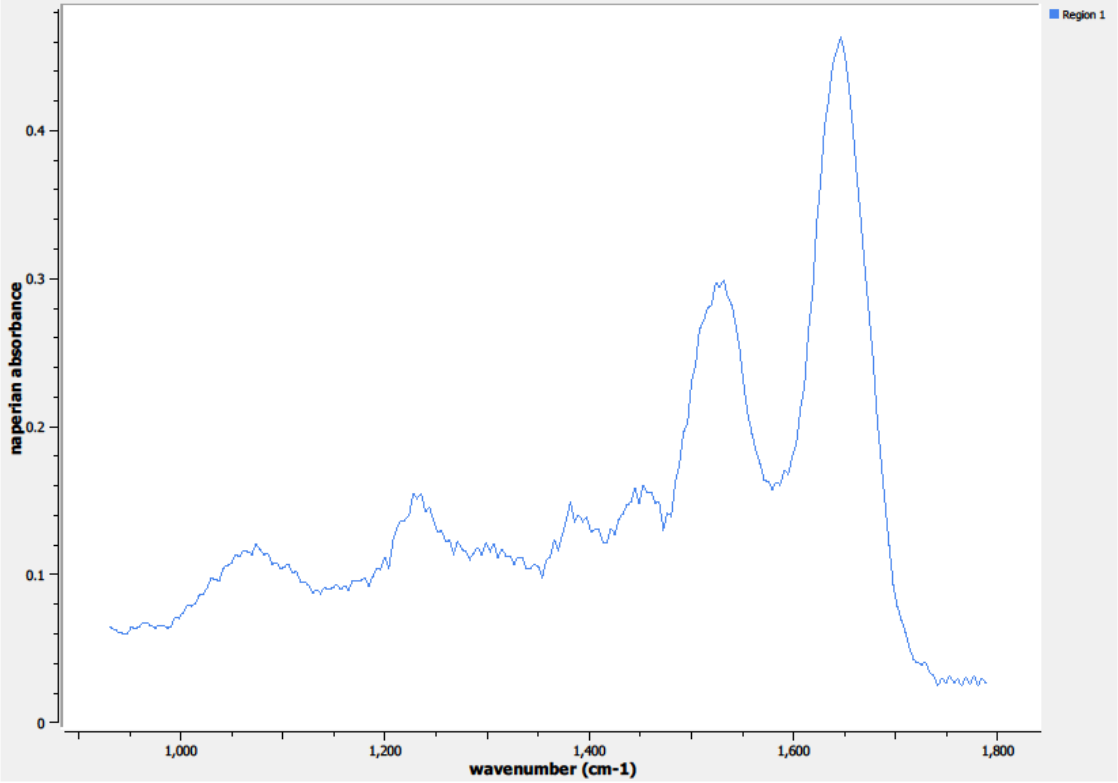



**Figure 0.20. Point U Spectra**

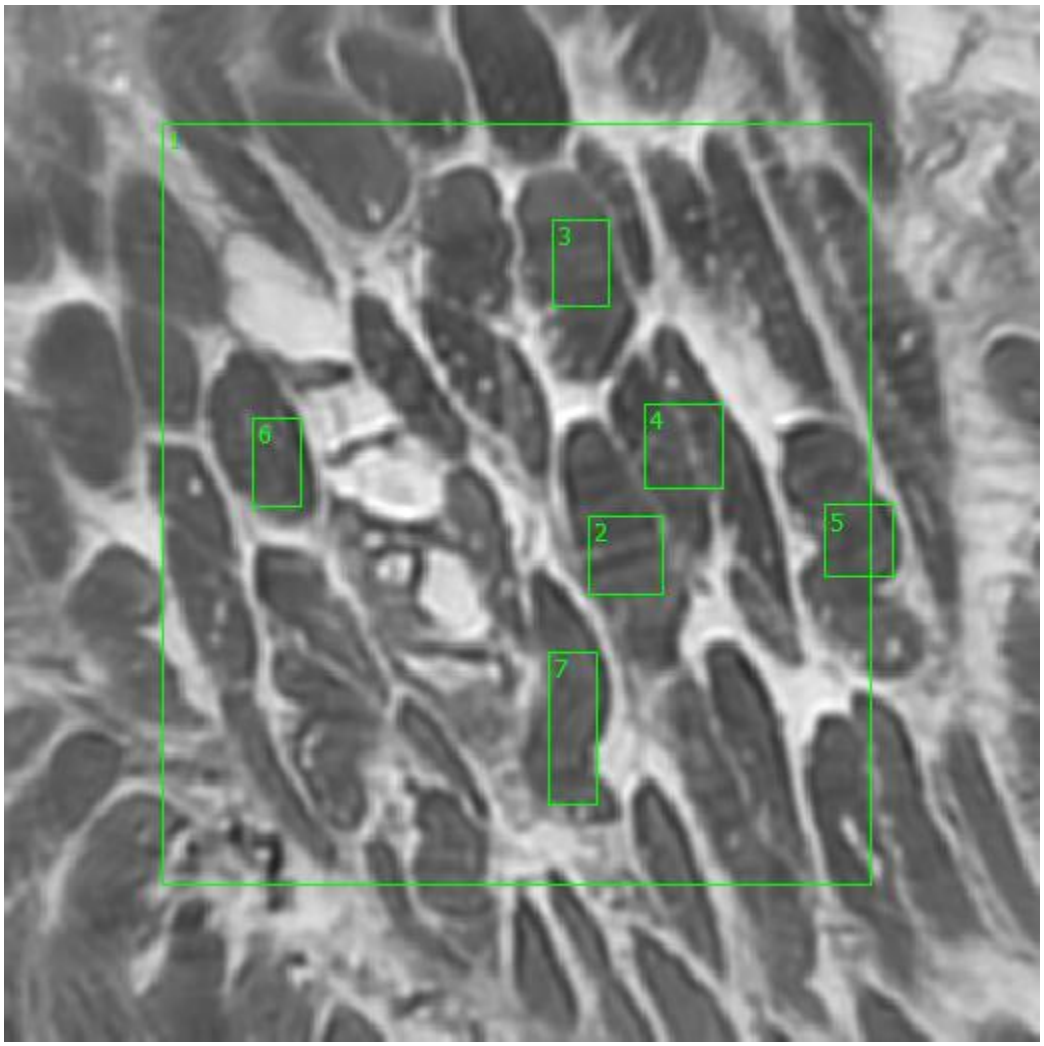



**Figure 0.21. Point V Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative, transition region**
**Regions 2, 3, 4, 5, 6, 7: breast, malignant ductal adenocarcinoma, triple negative, condensed transition region**

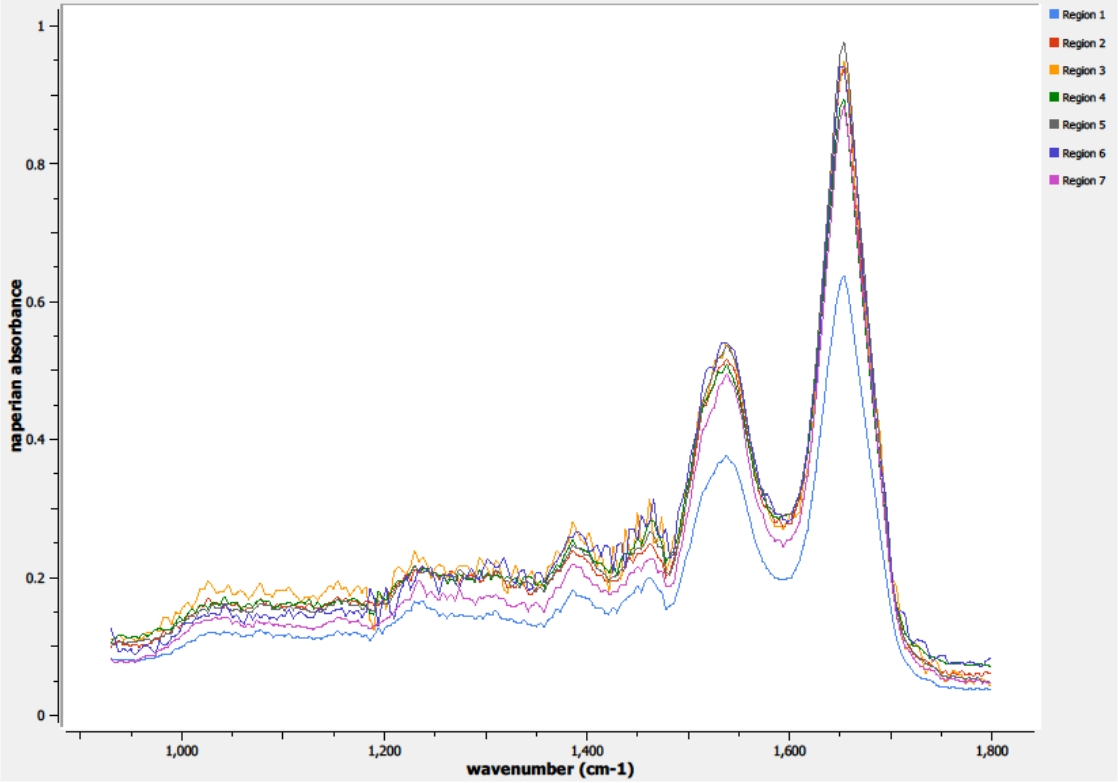



**Figure 0.22. Point V Spectra**

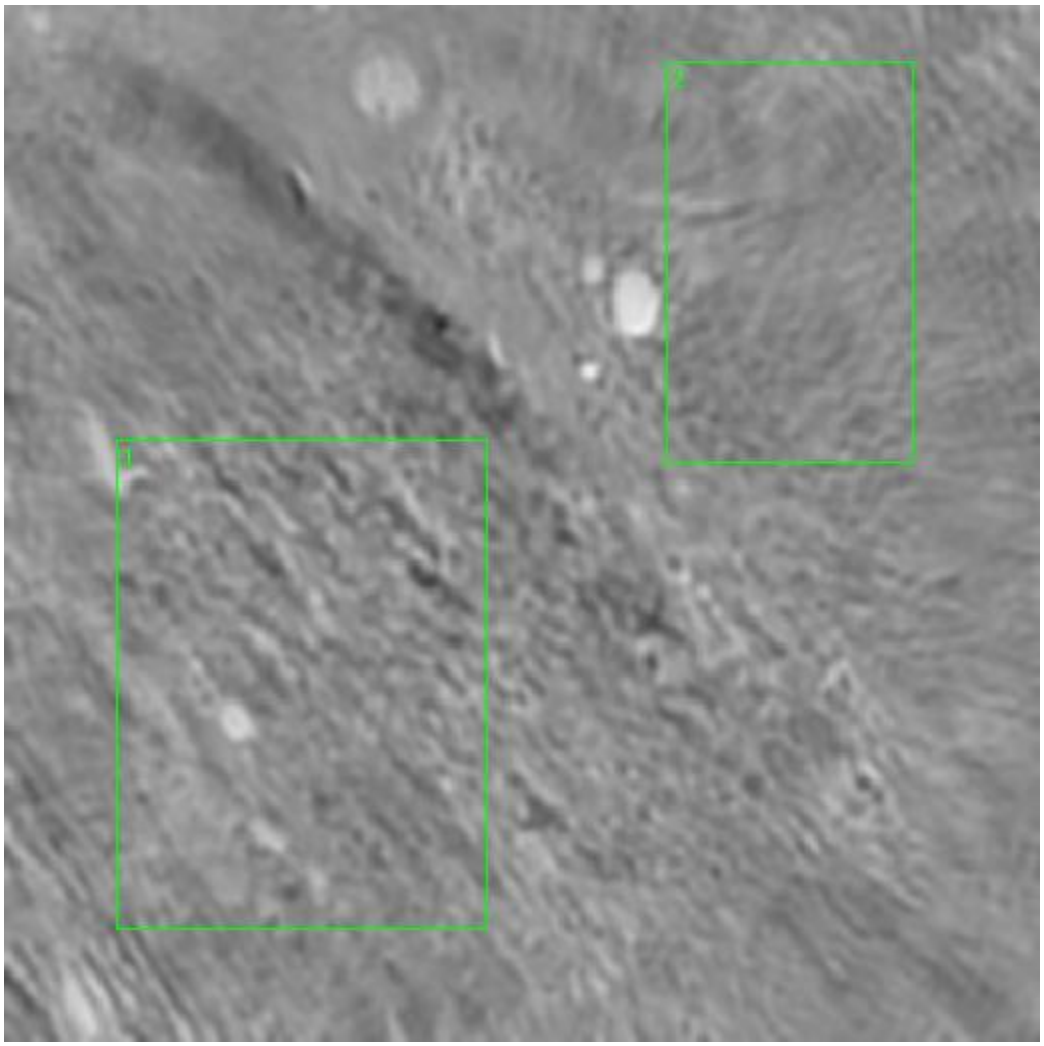


**Figure 0.23. Point W Regions of Interest**
**Regions 1, 2: breast, malignant ductal adenocarcinoma, triple negative**

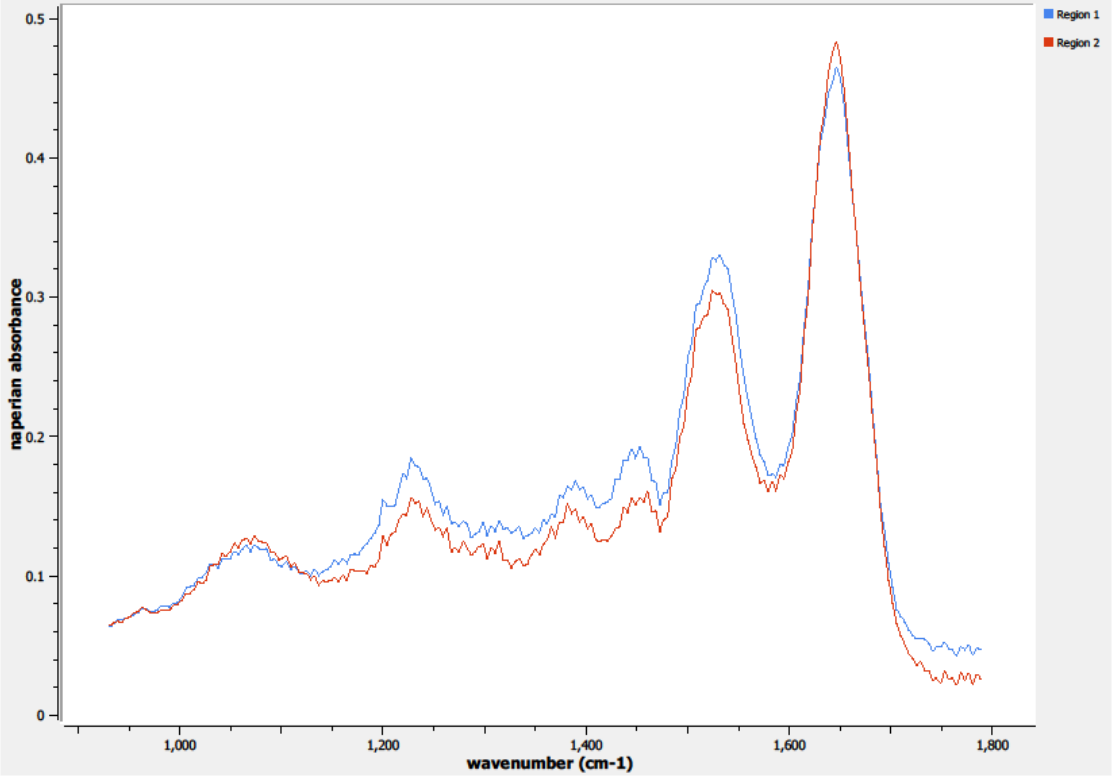



**Figure 0.24. Point W Spectra**

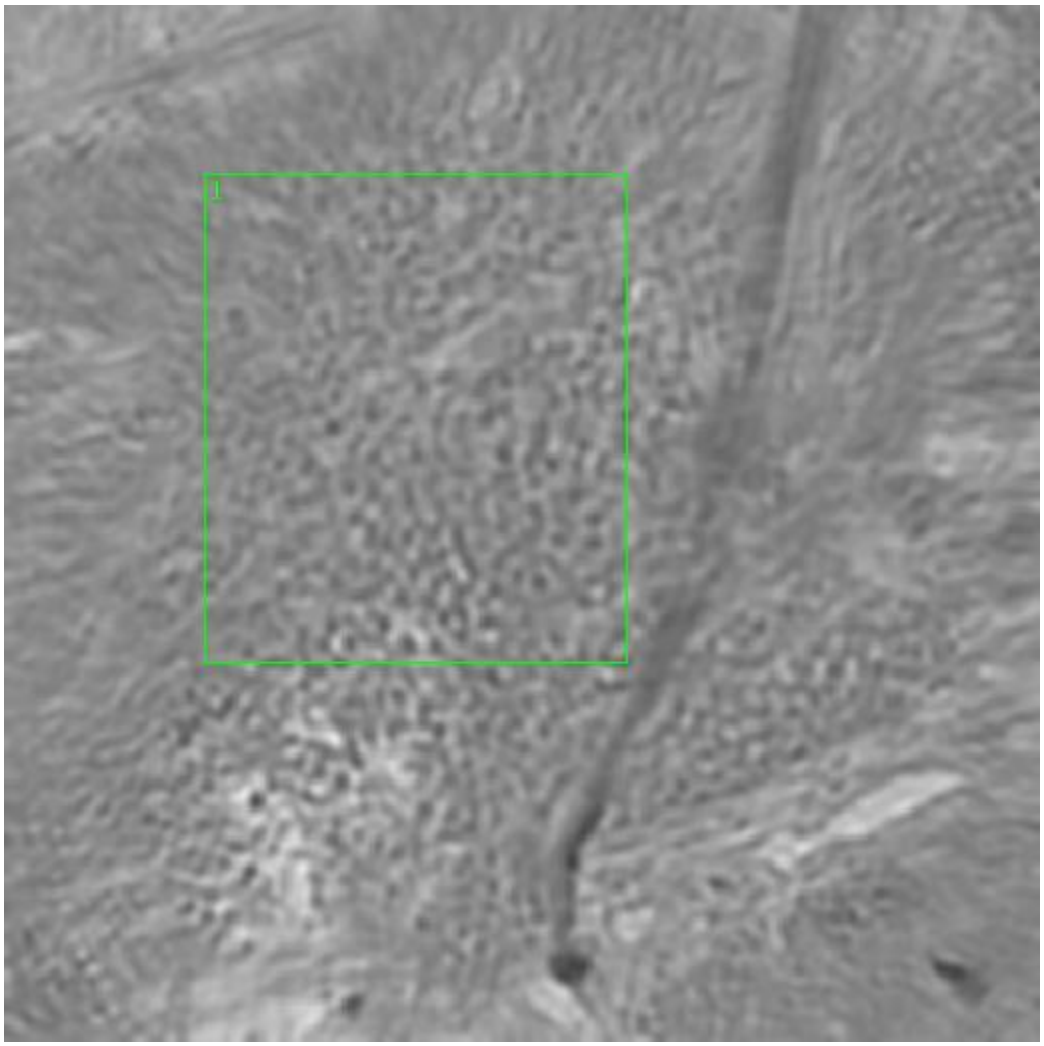


**Figure 0.25. Point X Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma, triple negative**

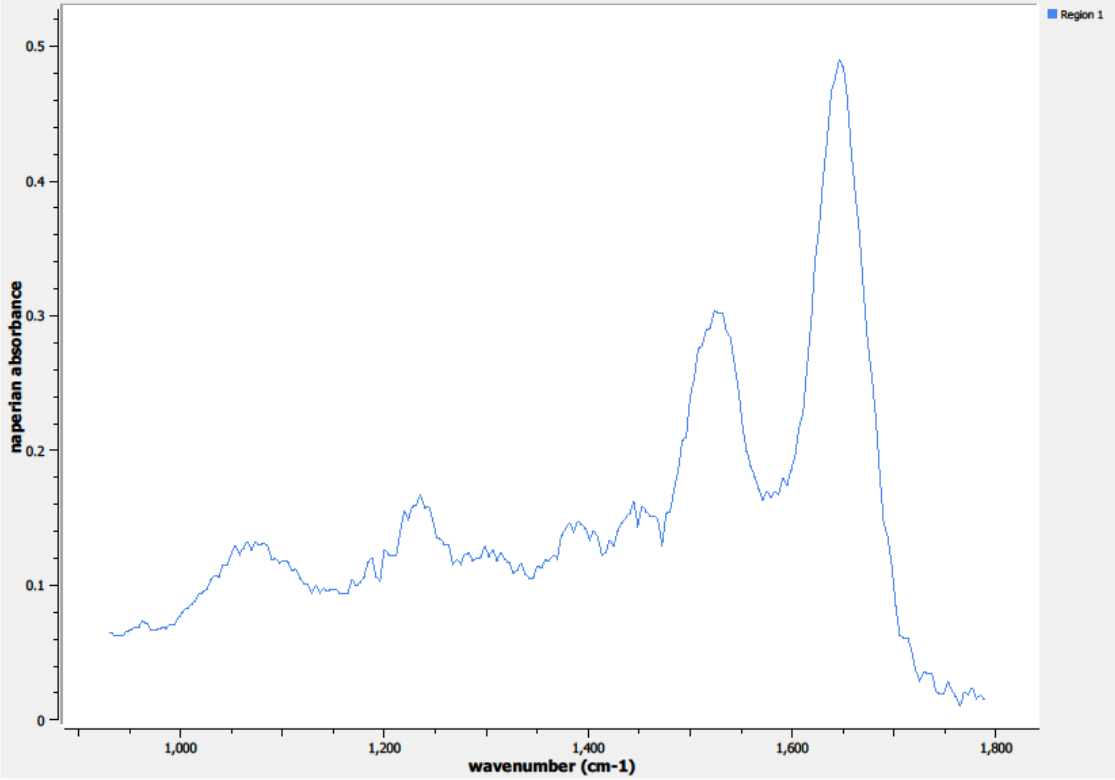



**Figure 0.26. Point X Spectra**

## A4.11-04-A001k

| | |
|---|---|
| Unique ID | 11-04-A001k |
| Age | 52 |
| Race | black |
| Location of Sample Collection | breast, ductal |
| Notes of Interest | malignant ductal adenocarcinoma, Stage 2, %T (80) / %N (20) |

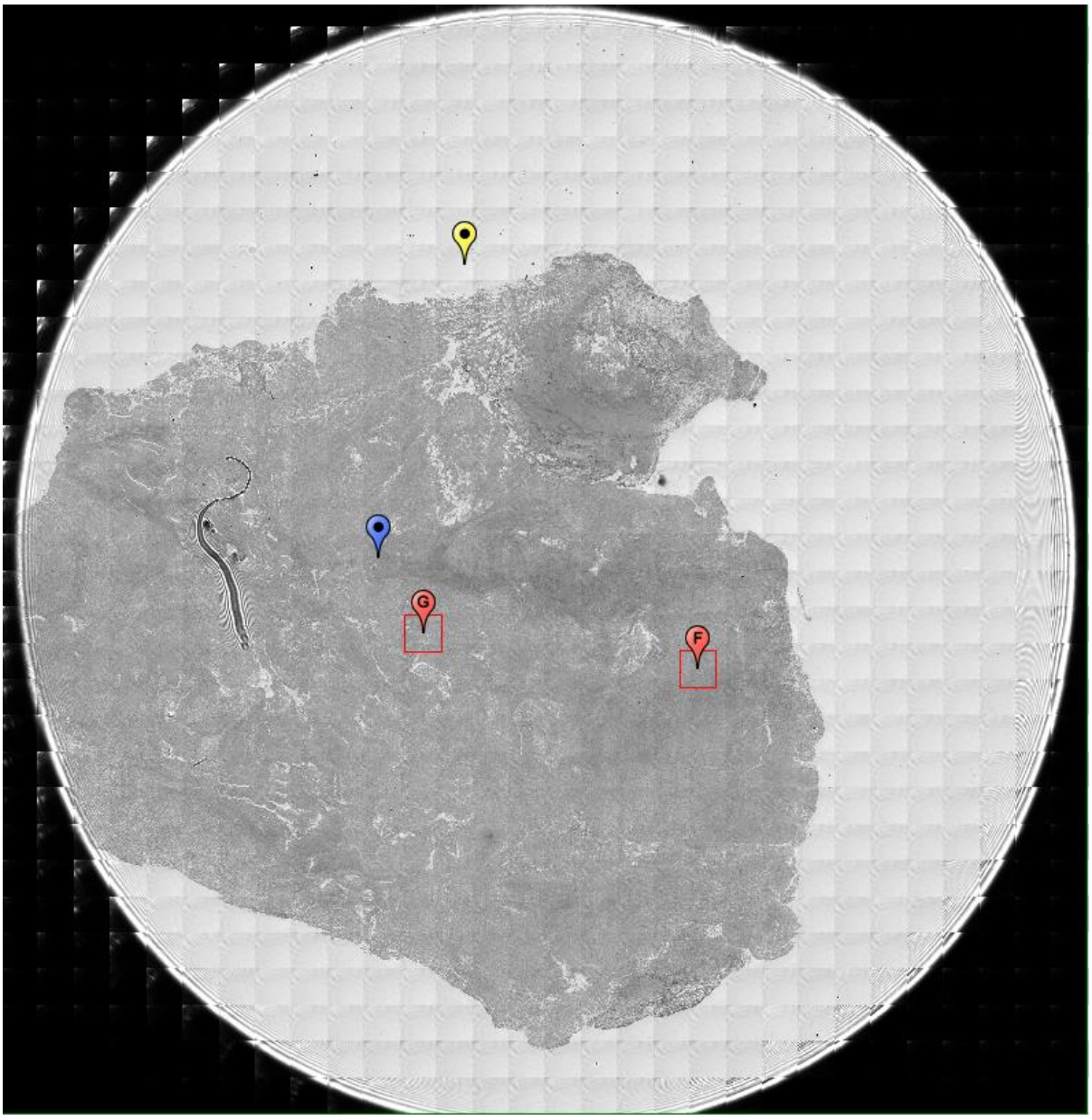



**Figure 0.27. Specimen 11-04-A001k, breast cancer tissue**

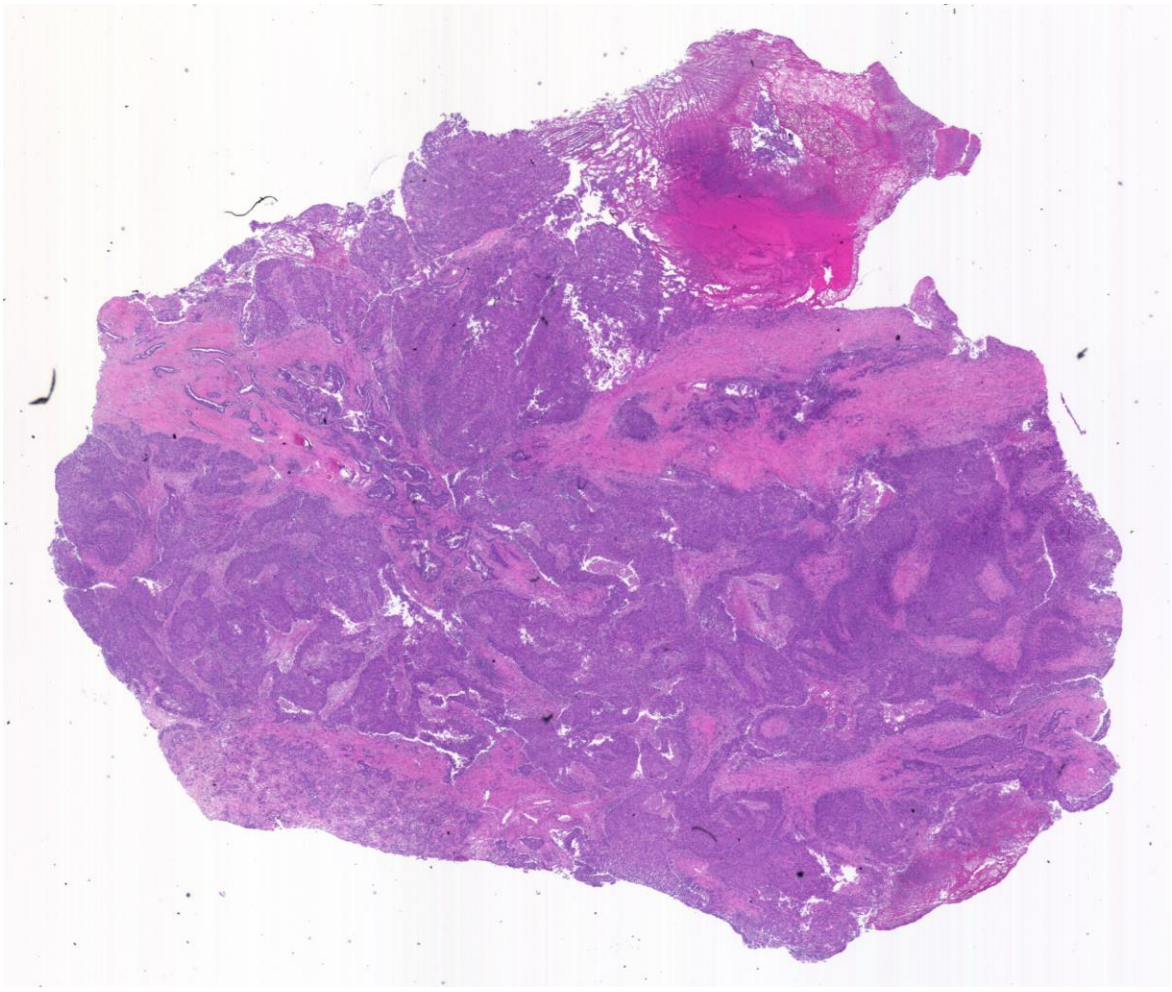


**Figure 0.28. Specimen 11-04-A001k, H&E stain, malignant ductal adenocarcinoma of the breast**

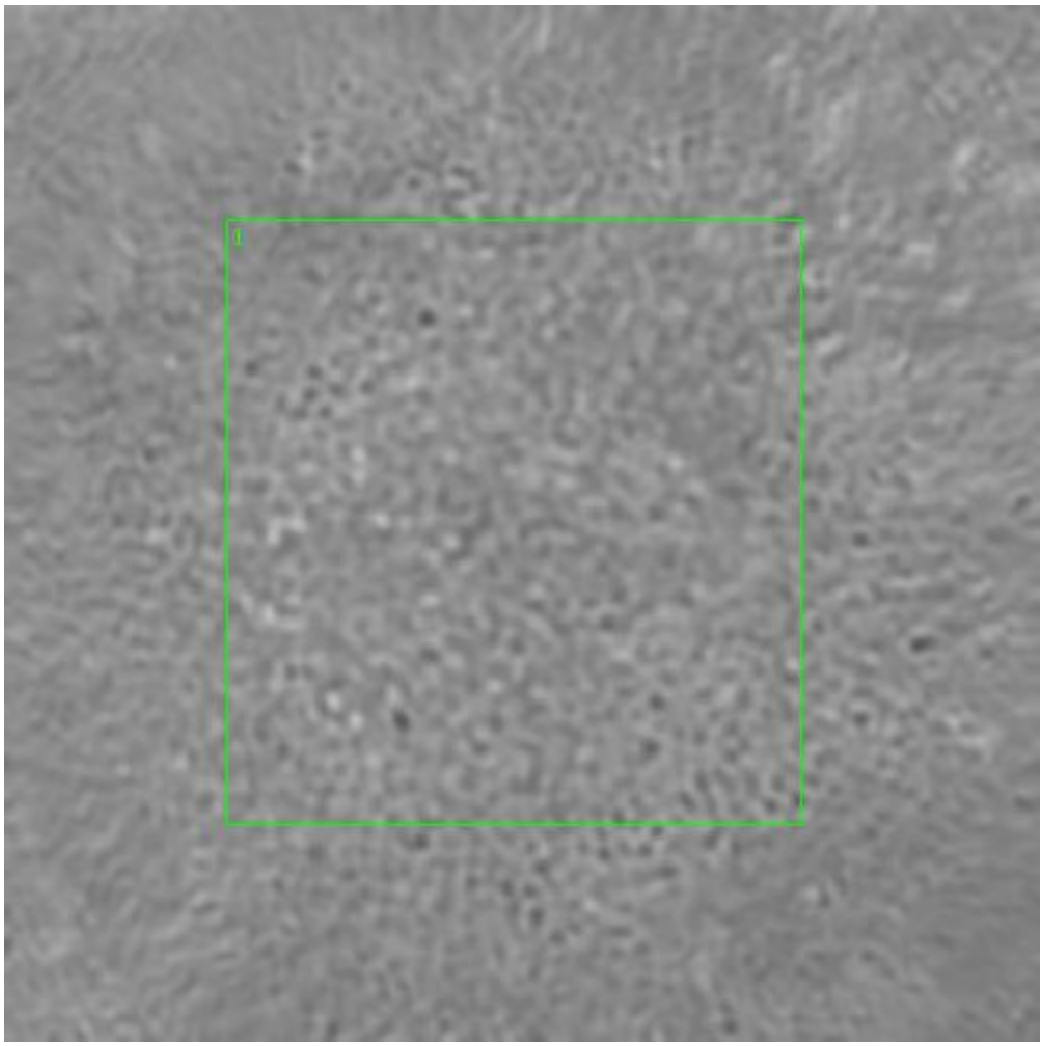



**Figure 0.29. Point F Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

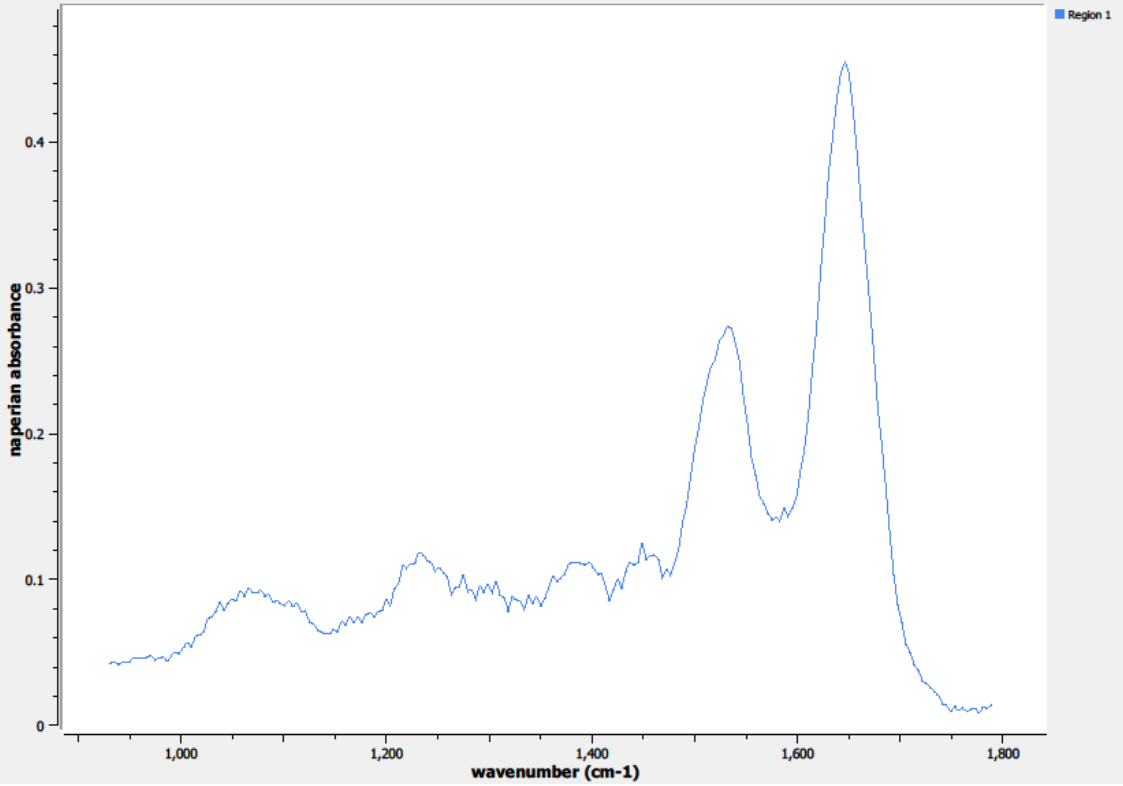



**Figure 0.30. Point F Spectra**

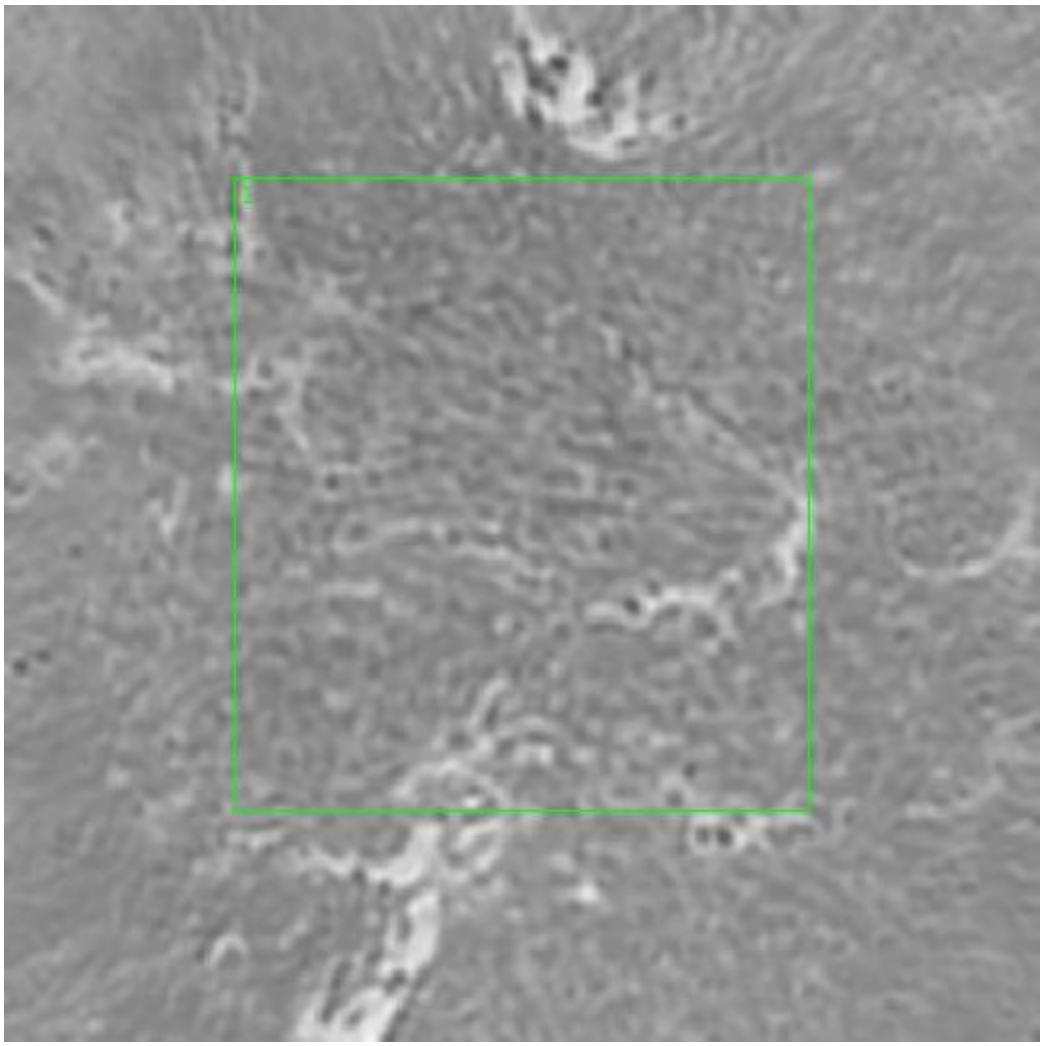


**Figure 0.31. Point G Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

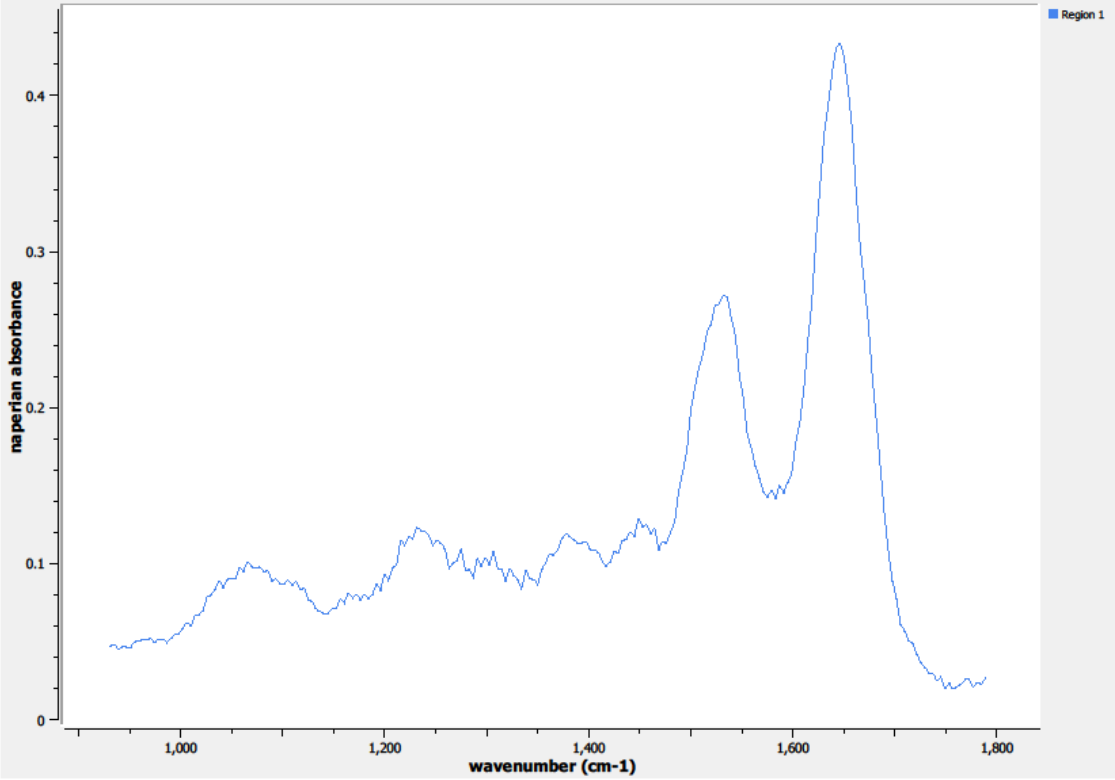



**Figure 0.32. Point G Spectra**

## A5.12-01-A021a-2

| | |
|---|---|
| Unique ID | 12-01-A021a-2 |
| Age | 49 |
| Race | white |
| Location of Sample Collection | right breast, lobular |
| Notes of Interest | malignant lobular adenocarcinoma, %T (50) / %N (50) |

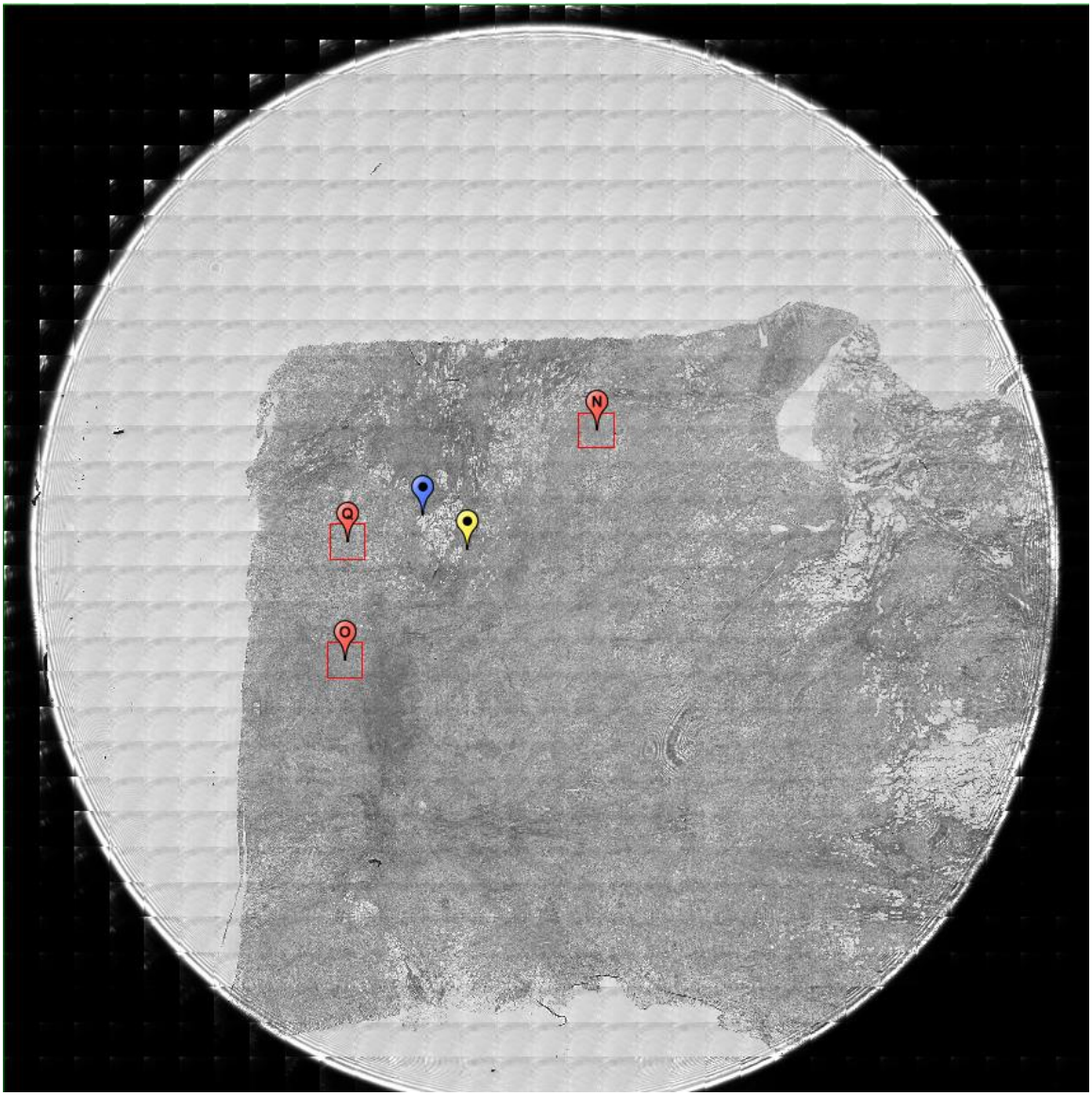



**Figure 0.33. Specimen 12-01-A021a-2, breast cancer tissue**

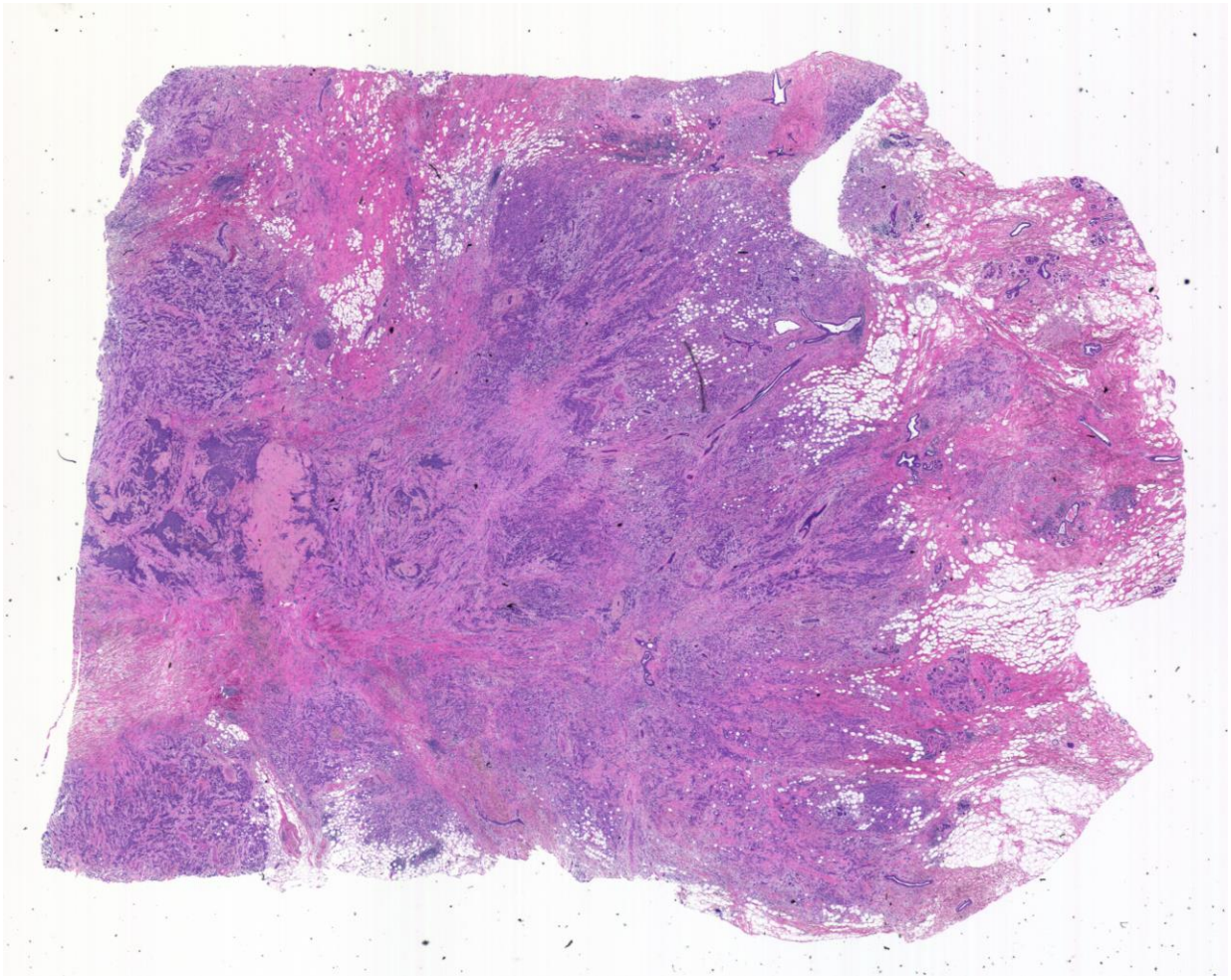


**Figure 0.34. Specimen 12-01-A021a-2, H&E stain, malignant lobular adenocarcinoma of the breast**

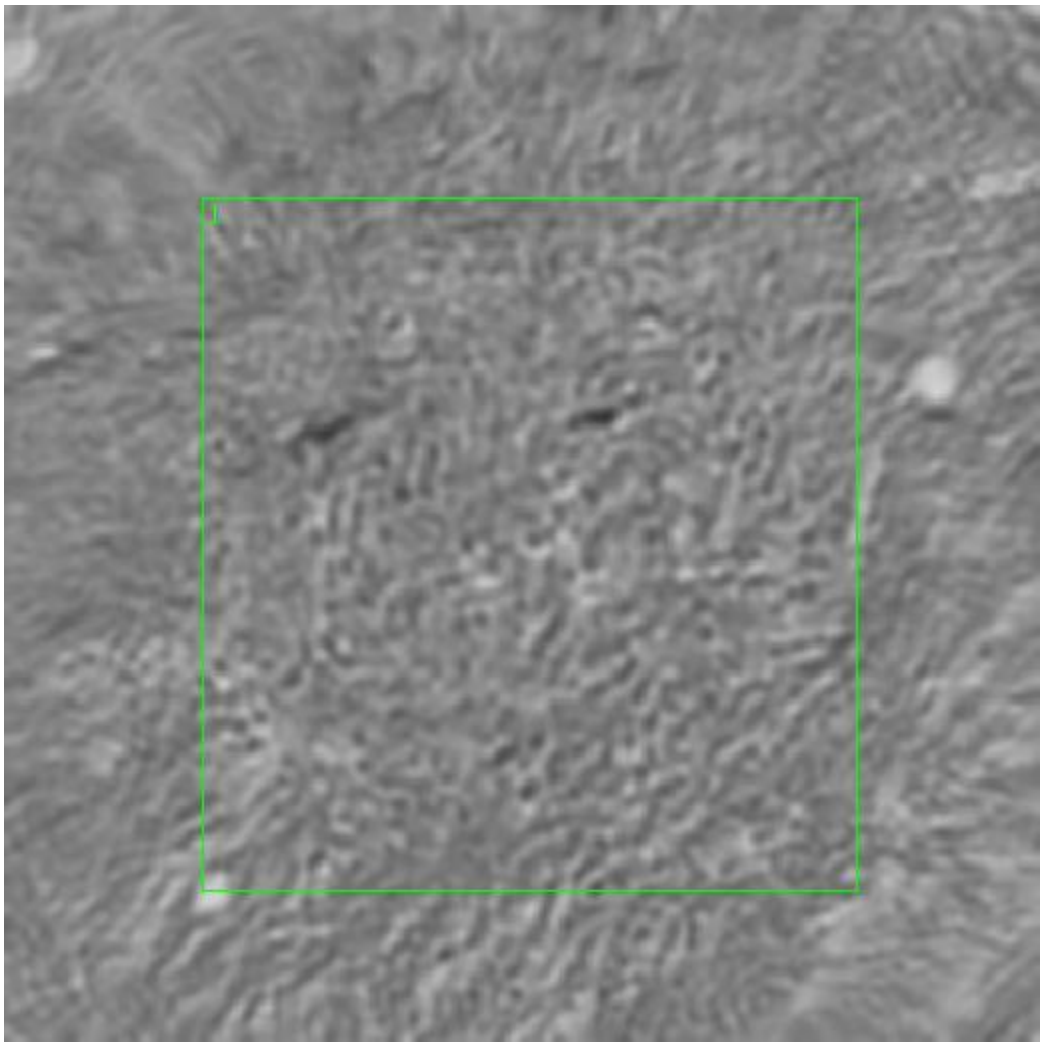


**Figure 0.35. Point N Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

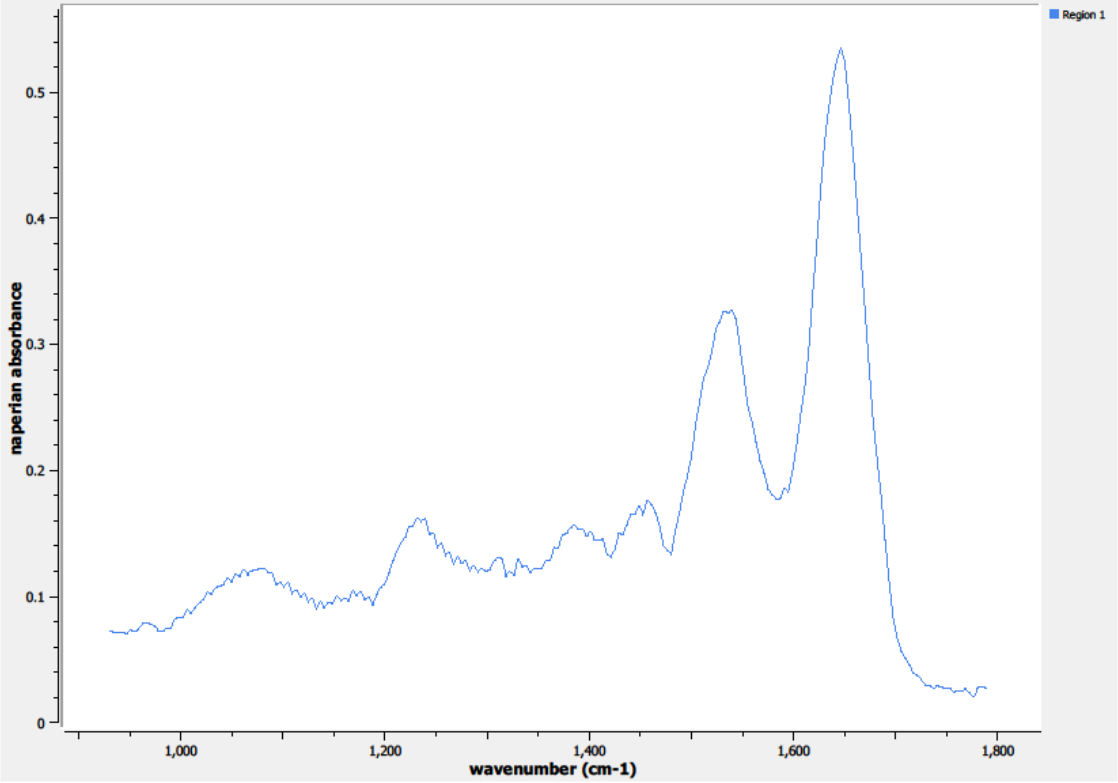



**Figure 0.36. Point N Spectra**

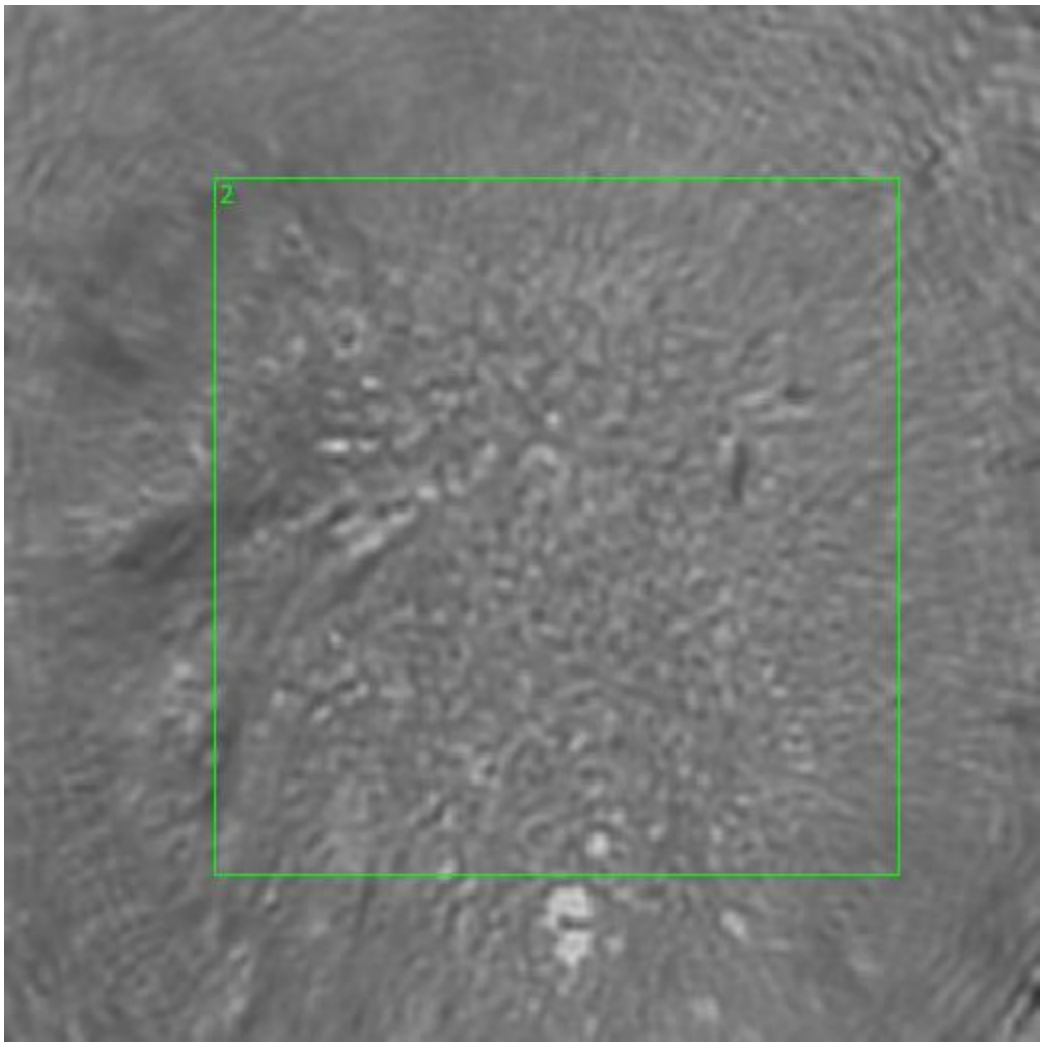



**Figure 0.37. Point O Regions of Interest**
**Region 2: breast, malignant lobular adenocarcinoma**

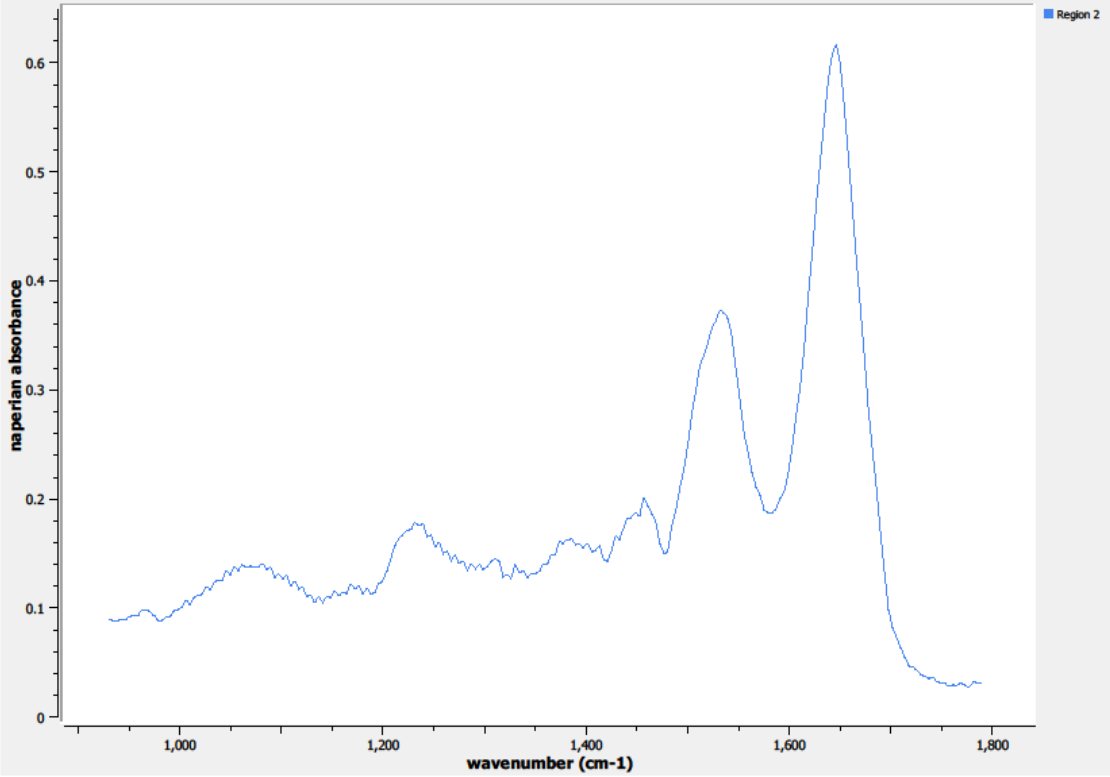



**Figure 0.38. Point O Spectra**

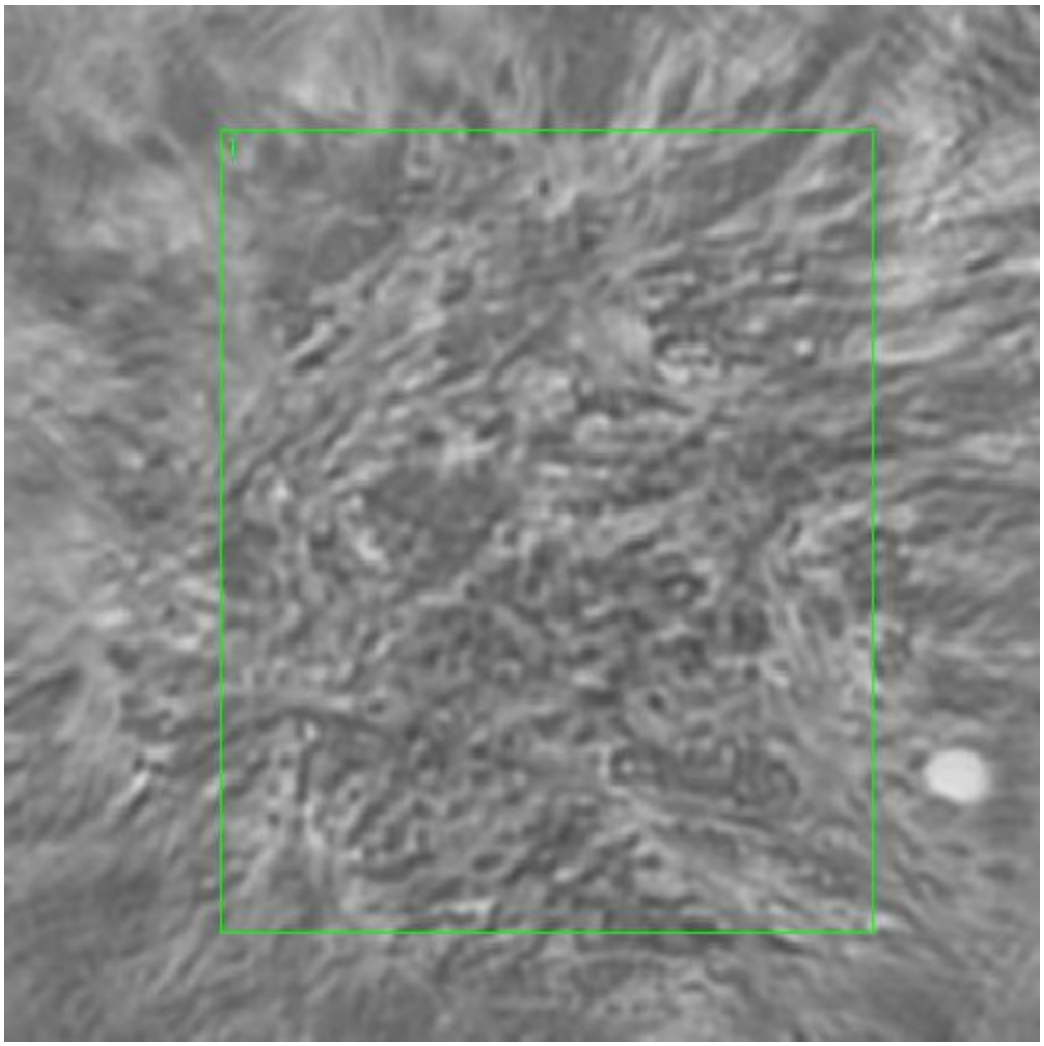


**Figure 0.39. Point Q Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

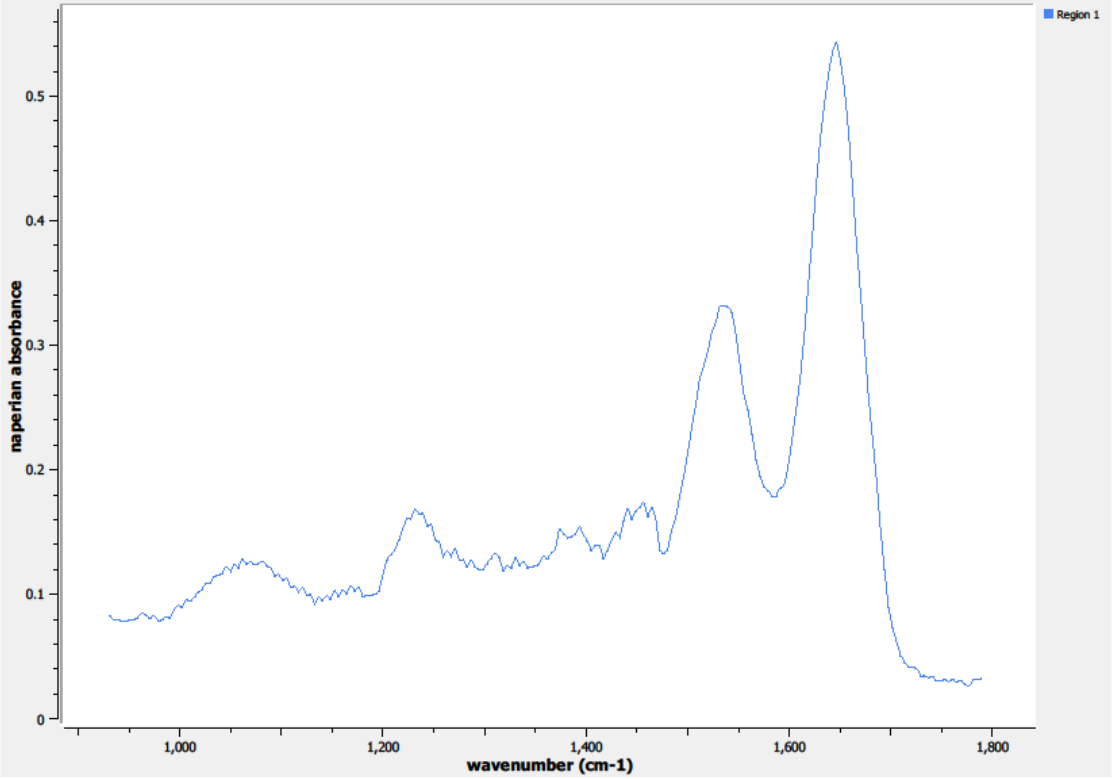



**Figure 0.40. Point Q Spectra**

## A6.12-04-A101f

| | |
|---|---|
| Unique ID | 12-04-A101f |
| Age | 62 |
| Race | white |
| Location of Sample Collection | right breast, lobular |
| Notes of Interest | malignant lobular adenocarcinoma, ER+, PR+, Her2/neu-, %T (70) / %N (70), may not use |

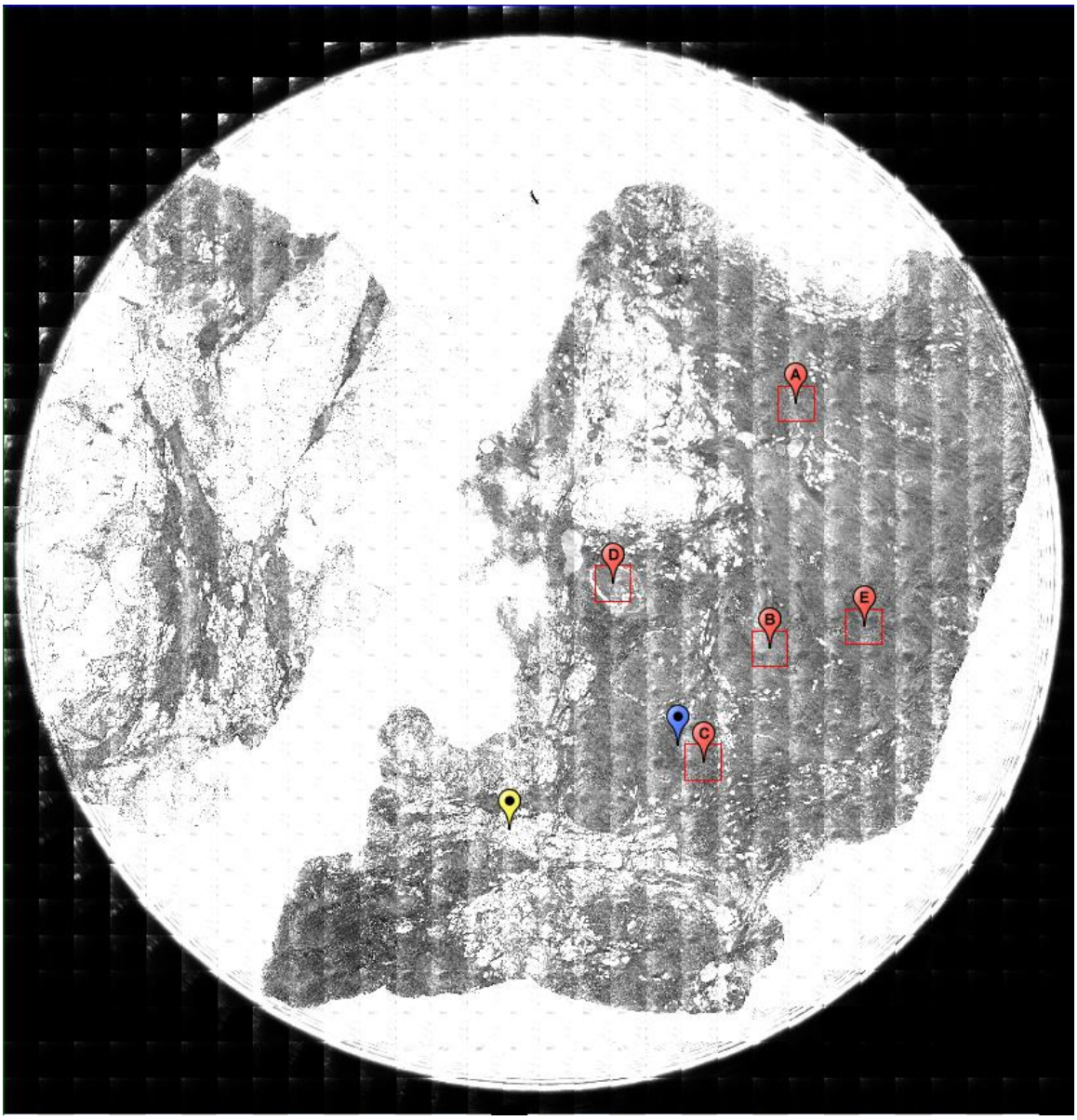



**Figure 0.41. Specimen 12-04-A101f, breast cancer tissue**

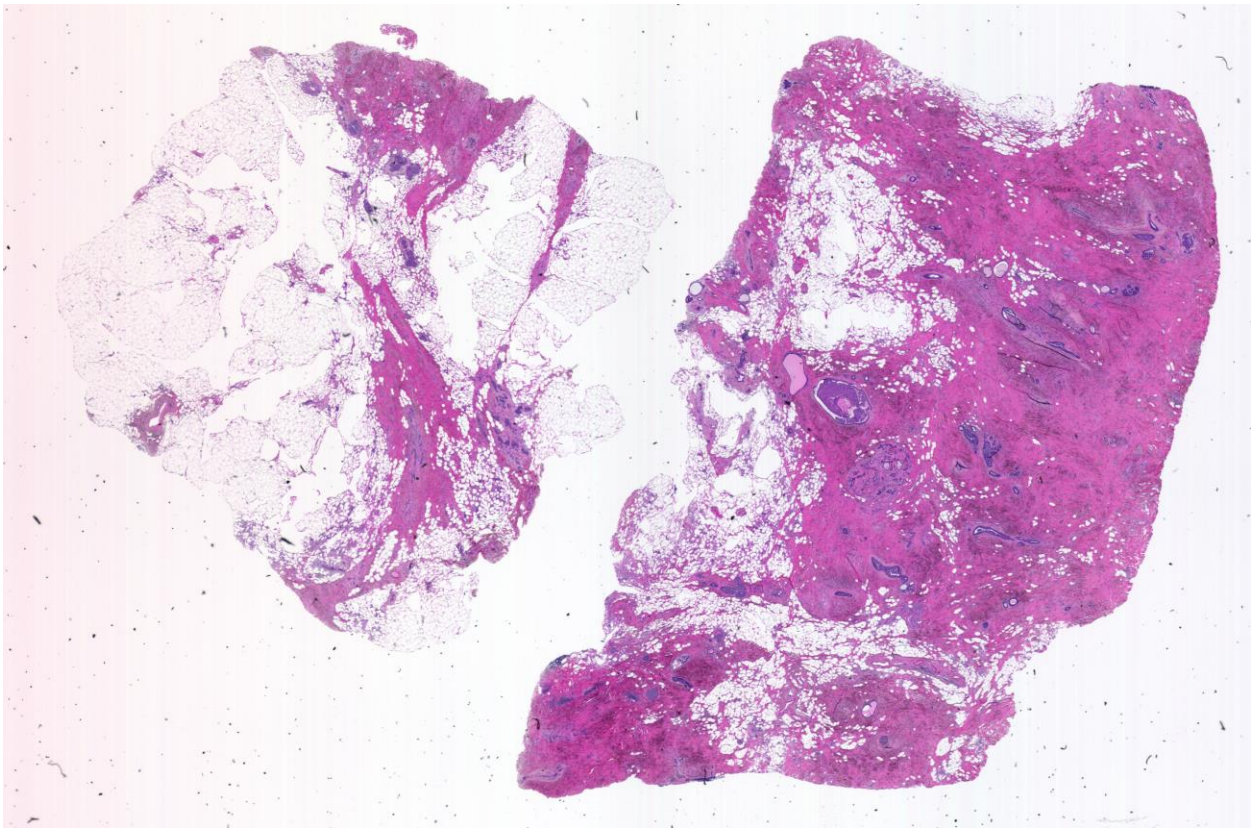


**Figure 0.42. Specimen 12-04-A101f, H&E stain, malignant lobular adenocarcinoma of the breast**

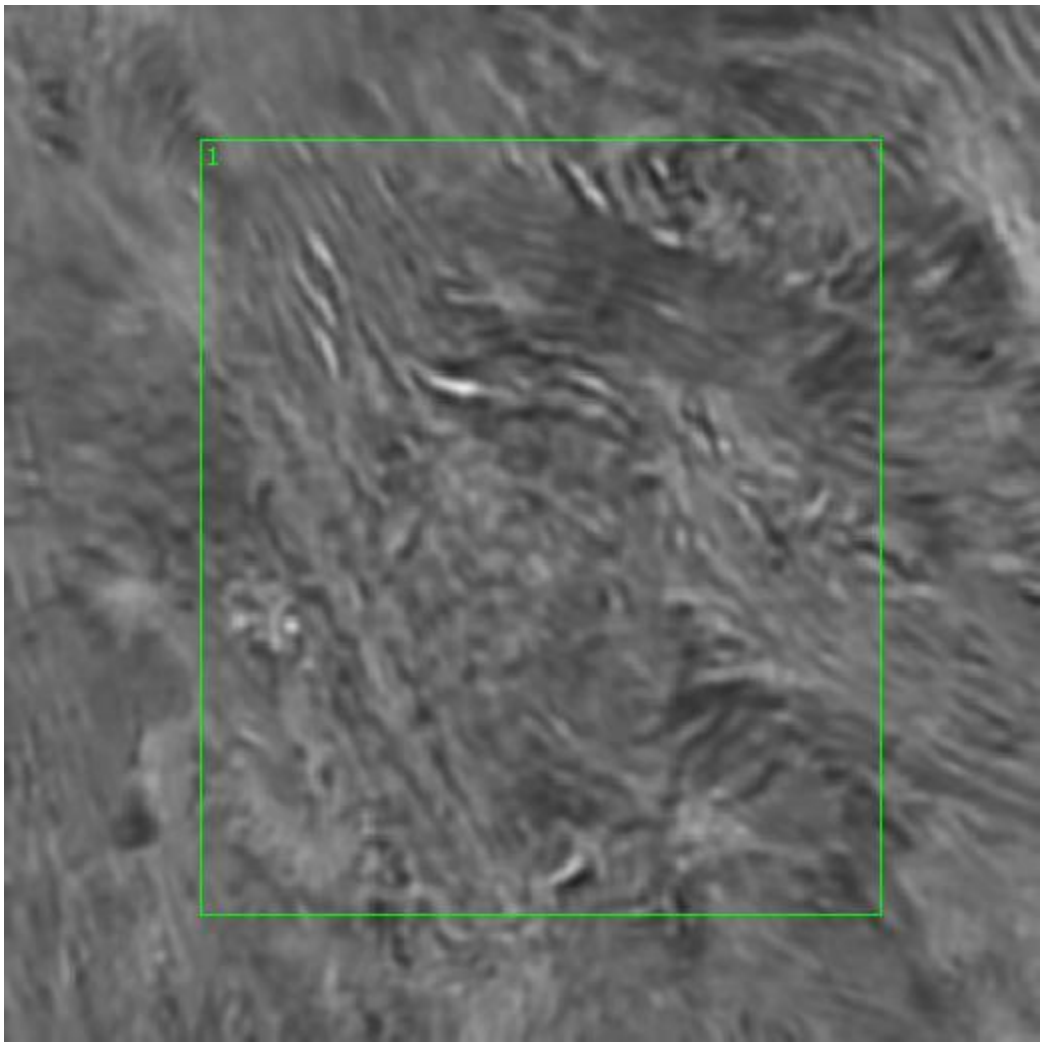



**Figure 0.43. Point A Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

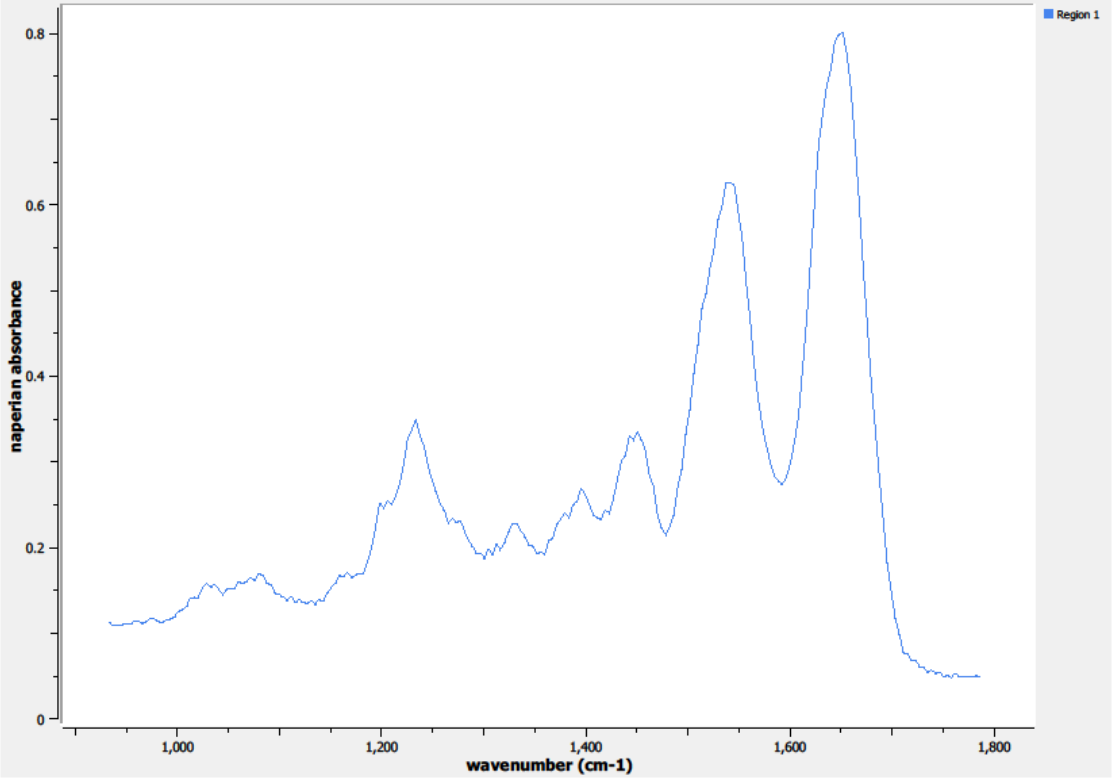



**Figure 0.44. Point A Spectra**

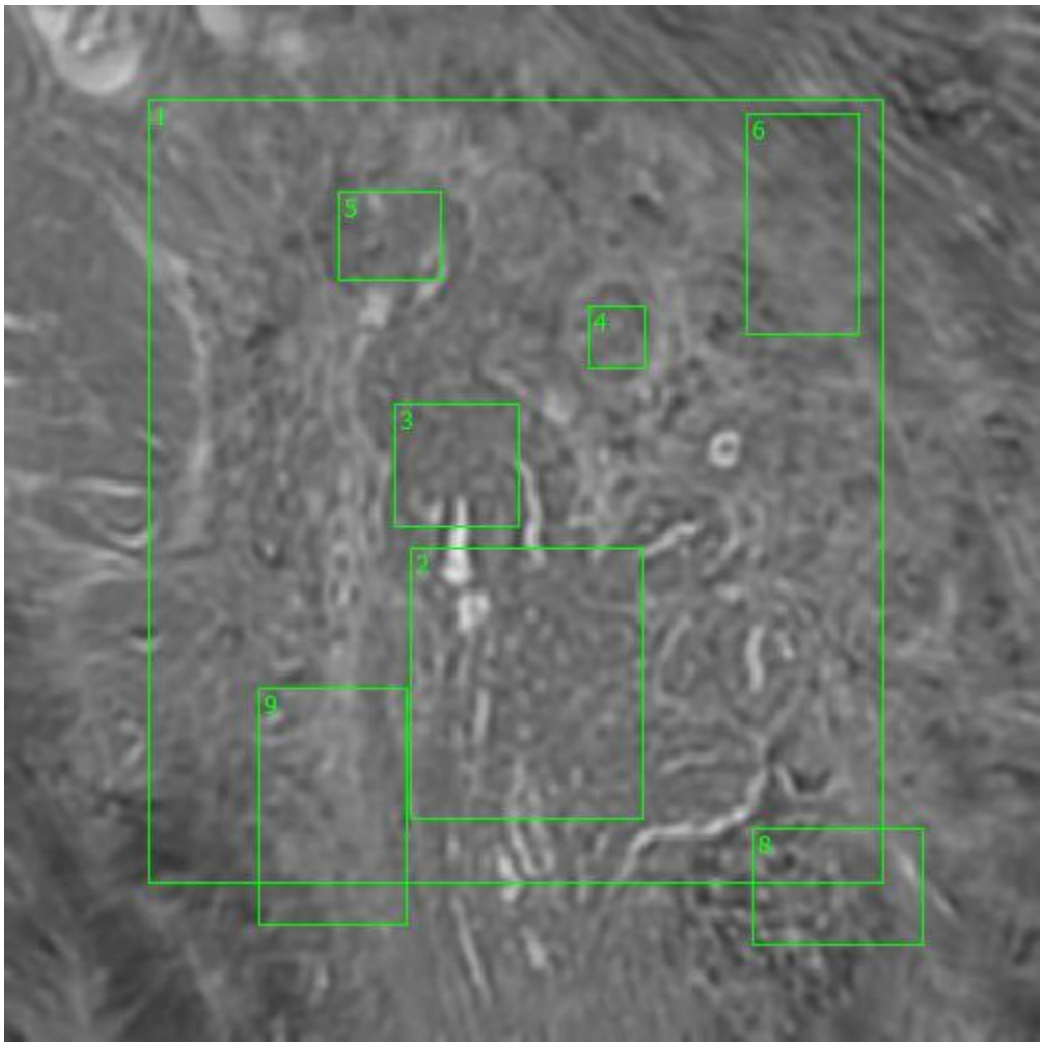



**Figure 0.45. Point B Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**
**Regions 2, 3, 4, 5, 6, 8, 9: breast, malignant lobular adenocarcinoma, concentrated region**

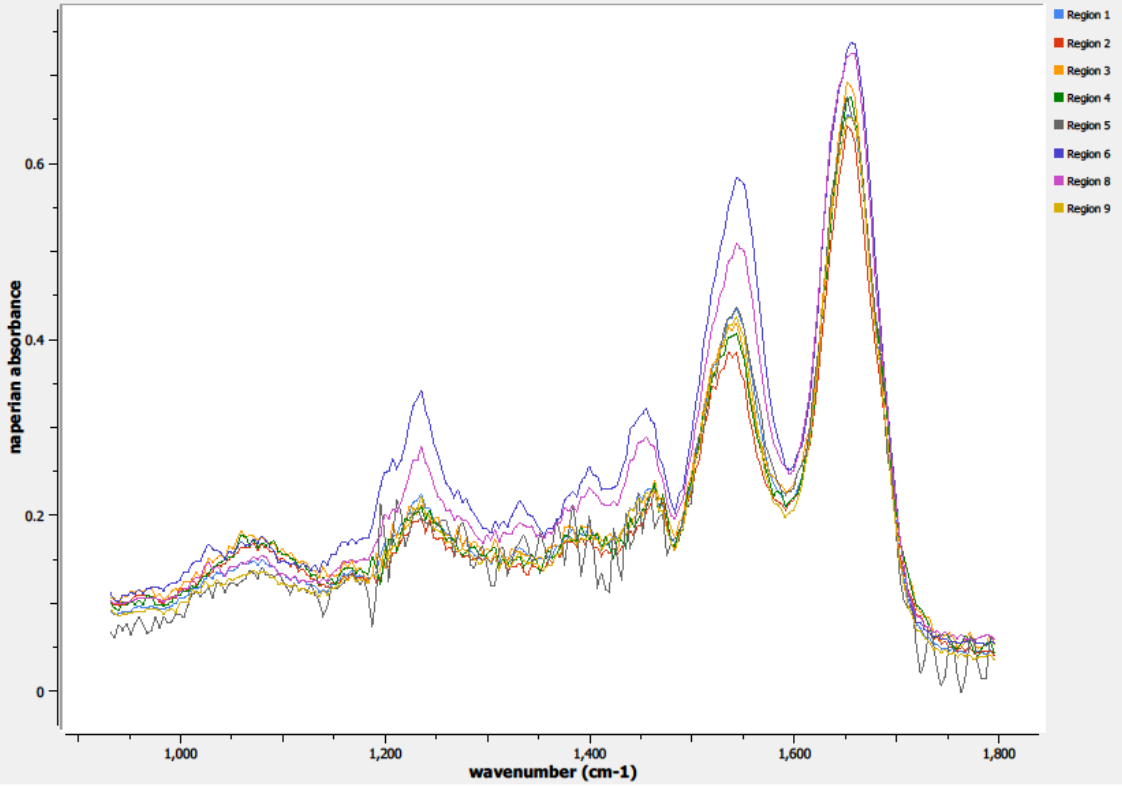



**Figure 0.46. Point B Spectra**

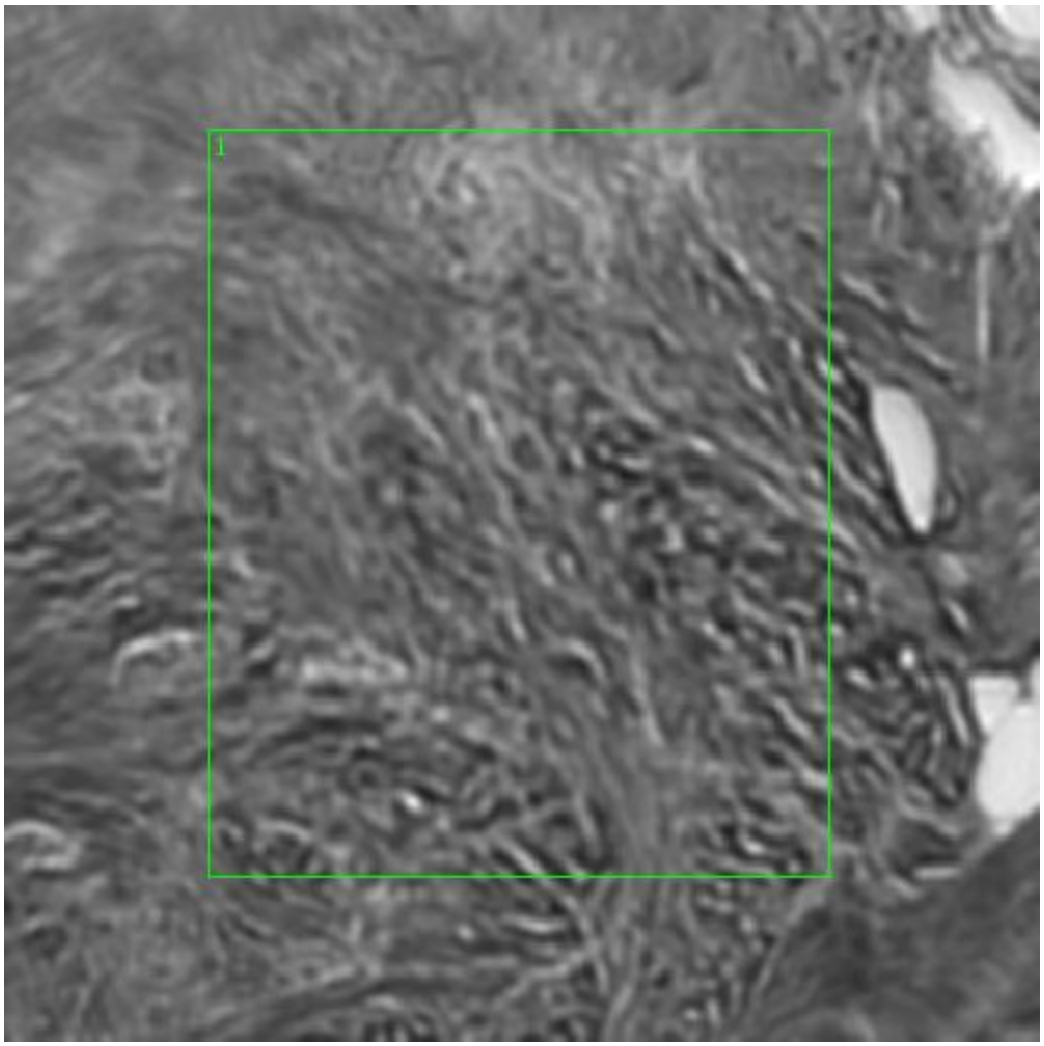



**Figure 0.47. Point C Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

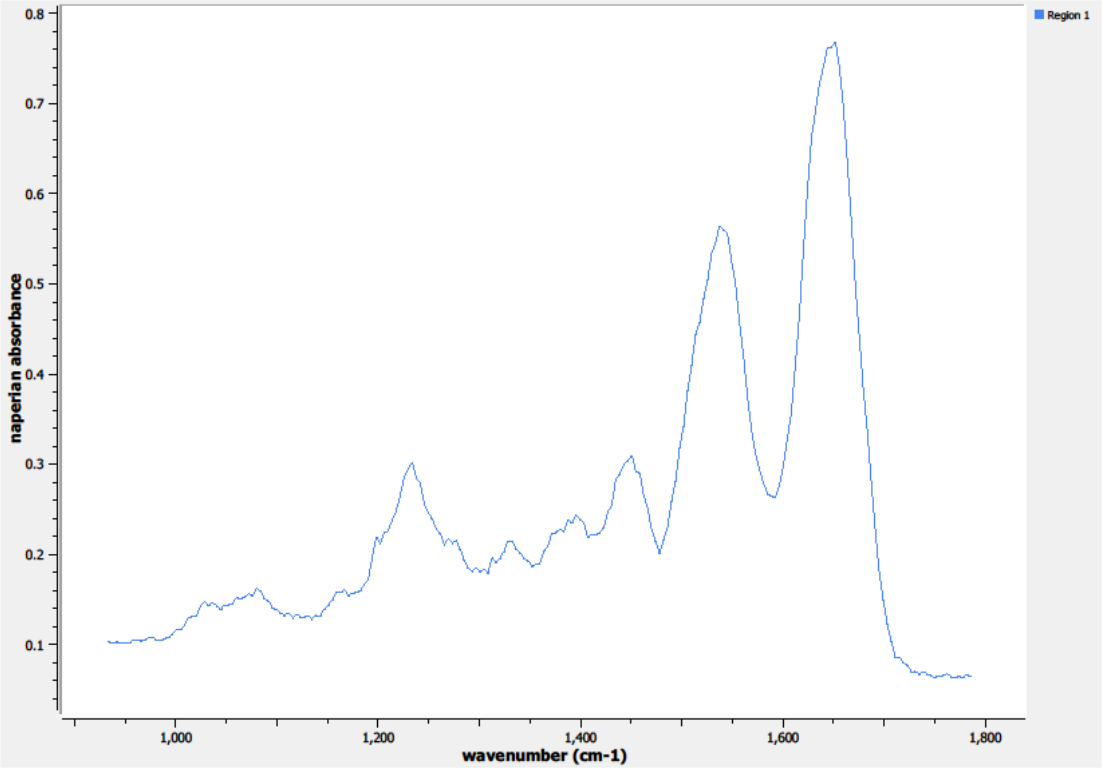



**Figure 0.48. Point C Spectra**

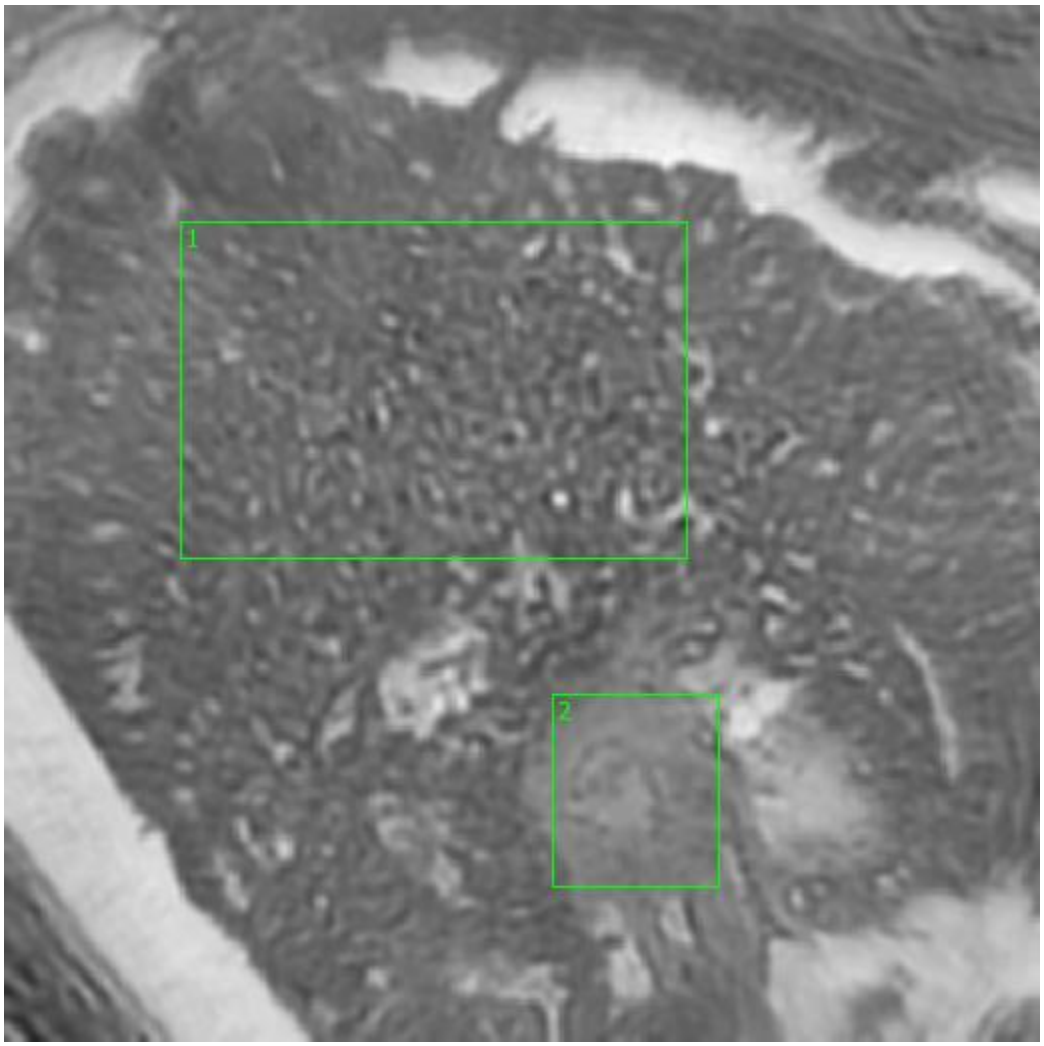



**Figure 0.49. Point D Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**
**Region 2: breast, malignant lobular adenocarcinoma, dense region**

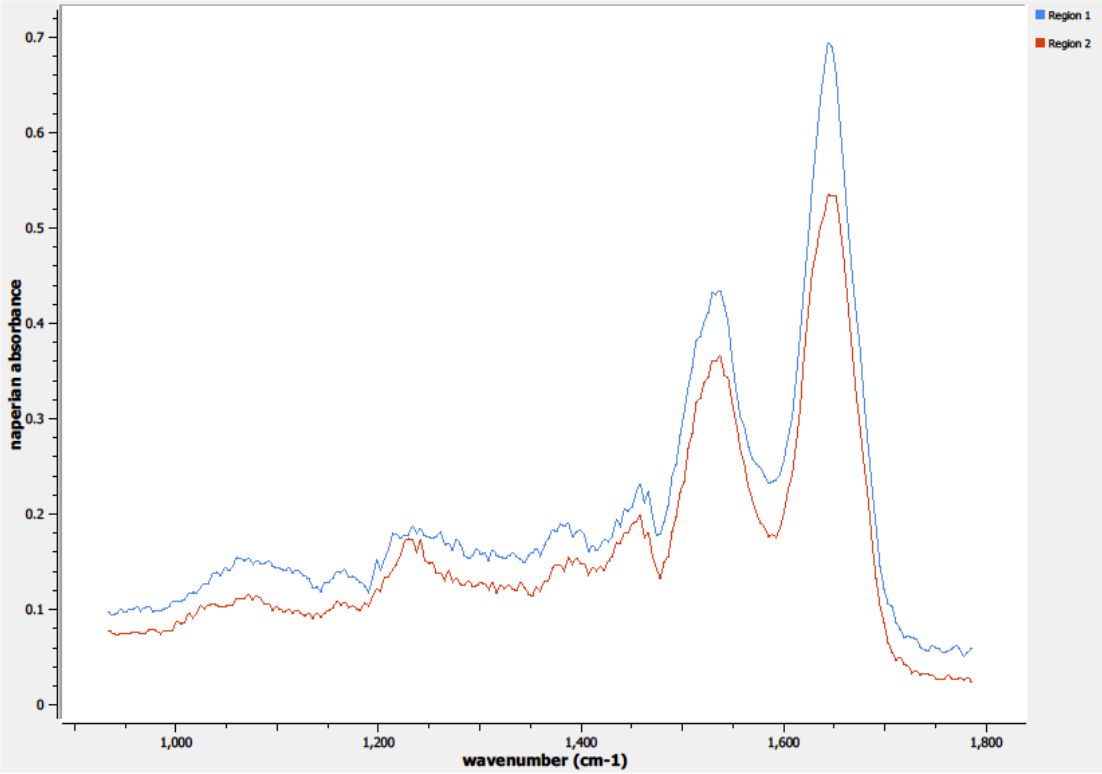



**Figure 0.50. Point D Spectra**

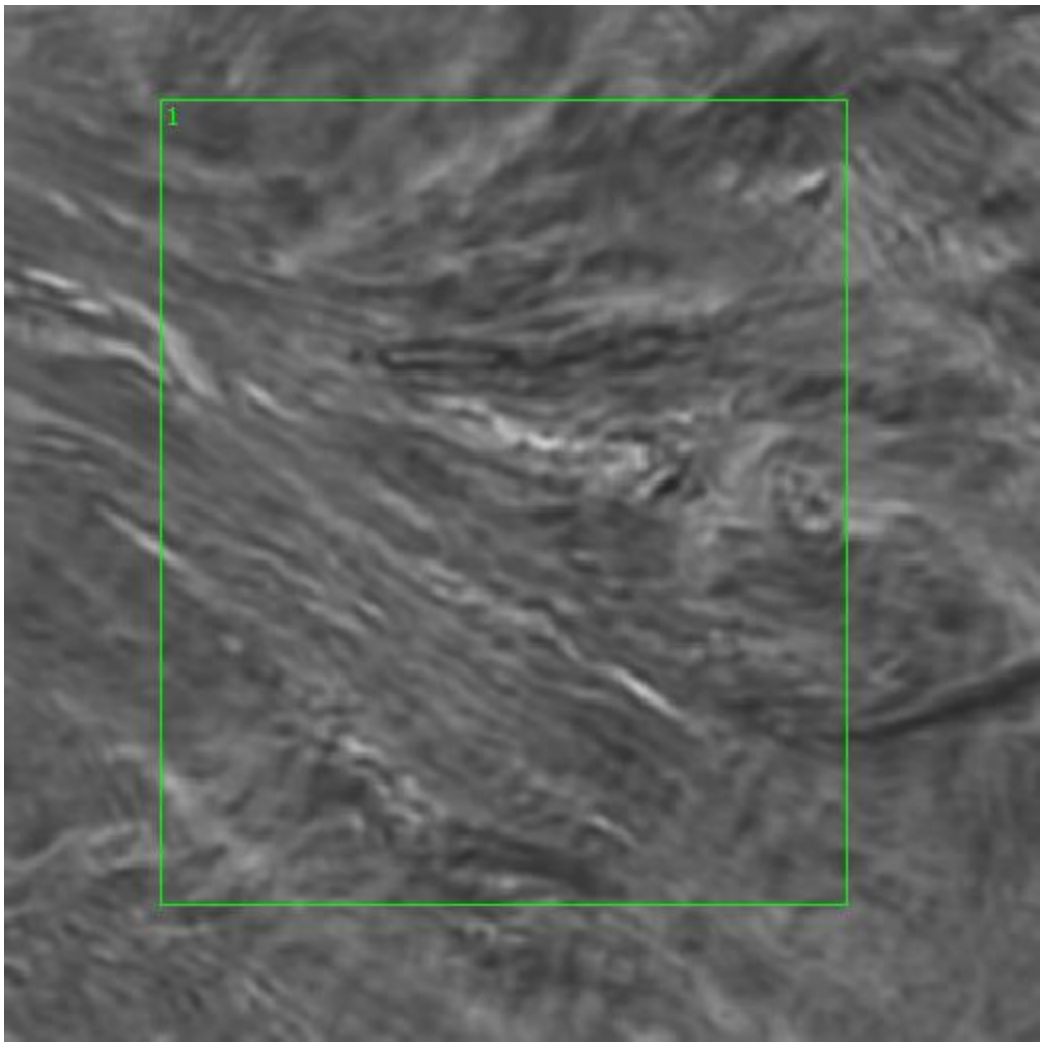



**Figure 0.51. Point E Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

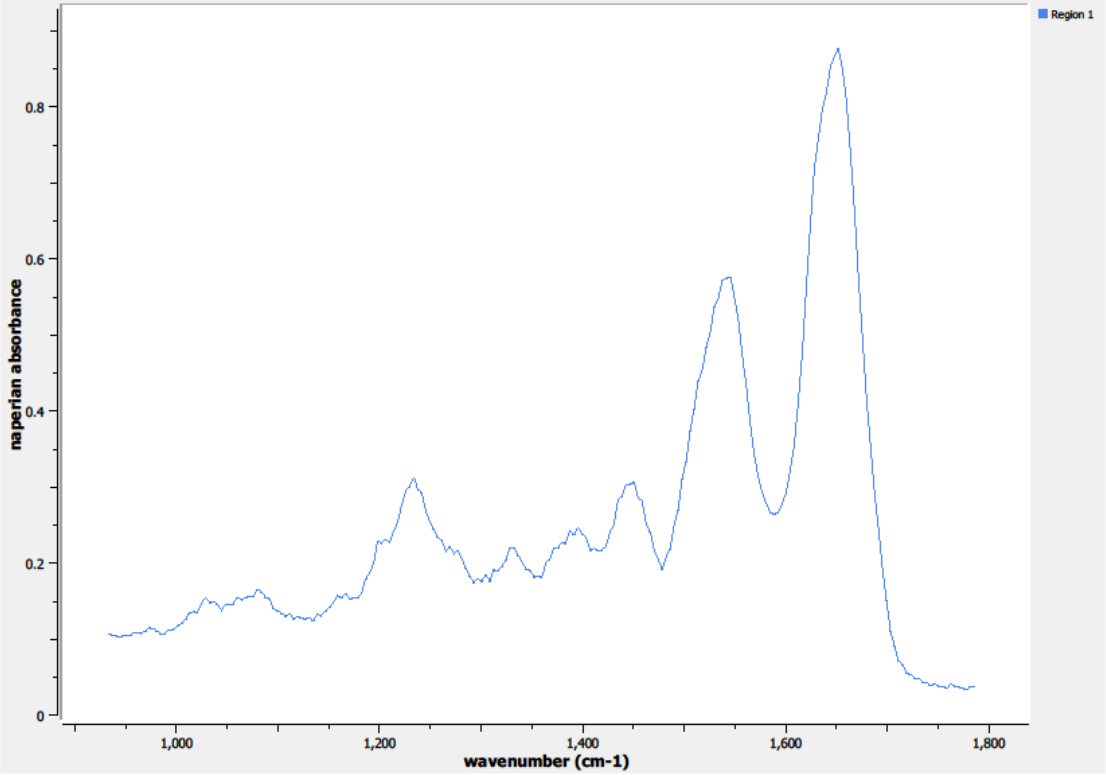



**Figure 0.52. Point E Spectra**

## A7.12-10-A032f

| | |
|---|---|
| Unique ID | 12-10-A032f |
| Age | 80 |
| Race | white |
| Location of Sample Collection | left breast, lobular |
| Notes of Interest | malignant lobular adenocarcinoma, T2N0M0, ER+, PR+, Her2/neu-, %T (90) / %N (30) |

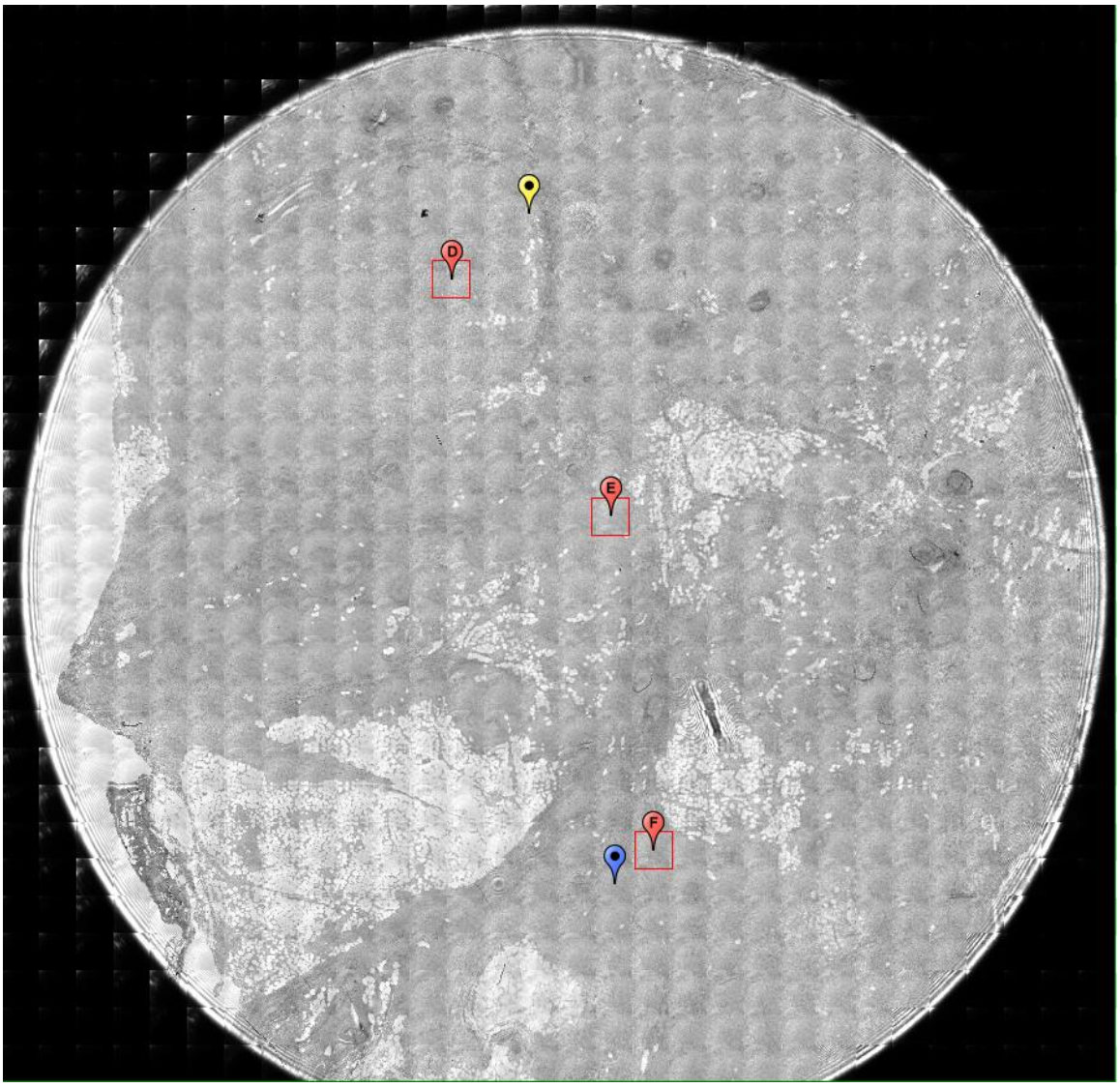



**Figure 0.53. Specimen 12-10-A032f, breast cancer tissue**

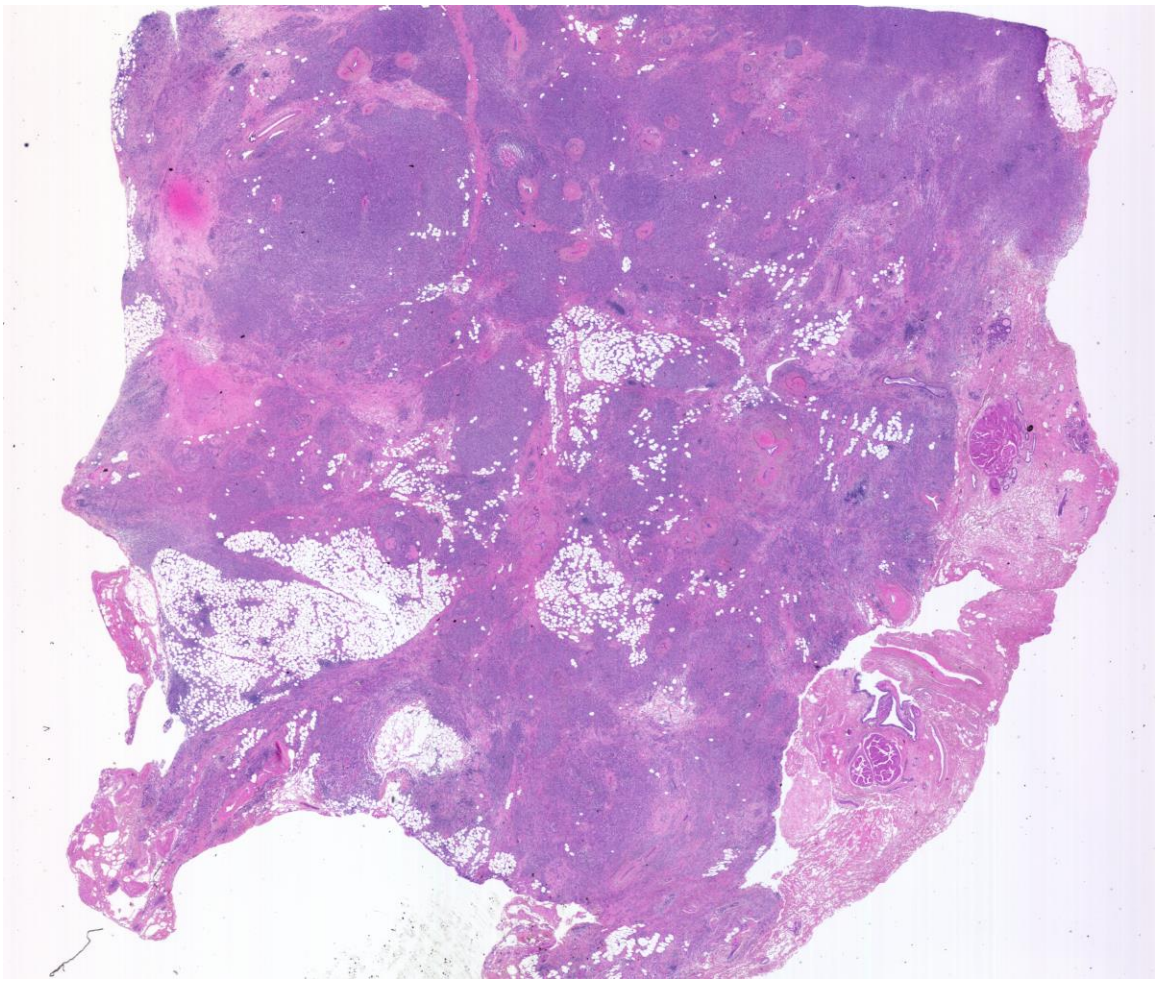


**Figure 0.54. Specimen 12-10-A032f, H&E stain, malignant lobular adenocarcinoma of the breast**

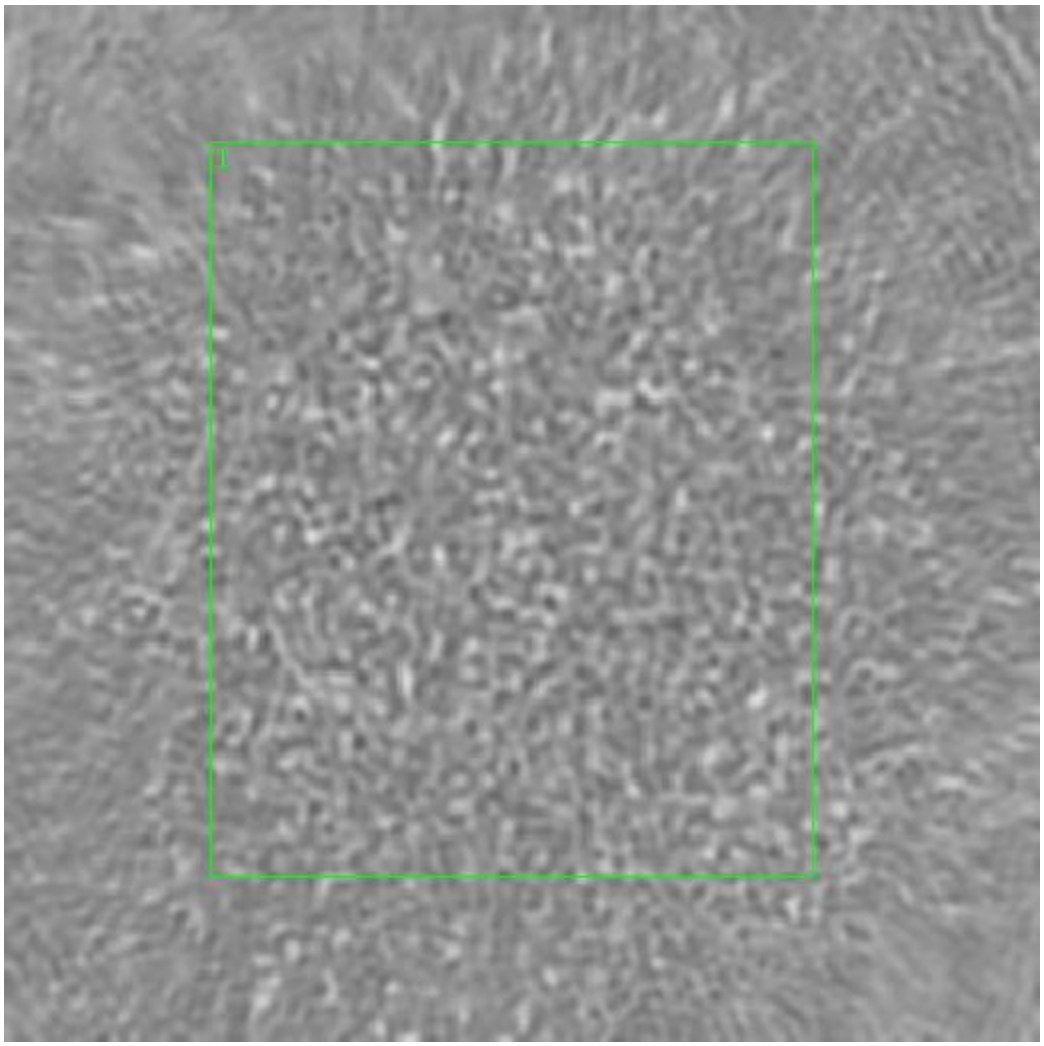


**Figure 0.55. Point D Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

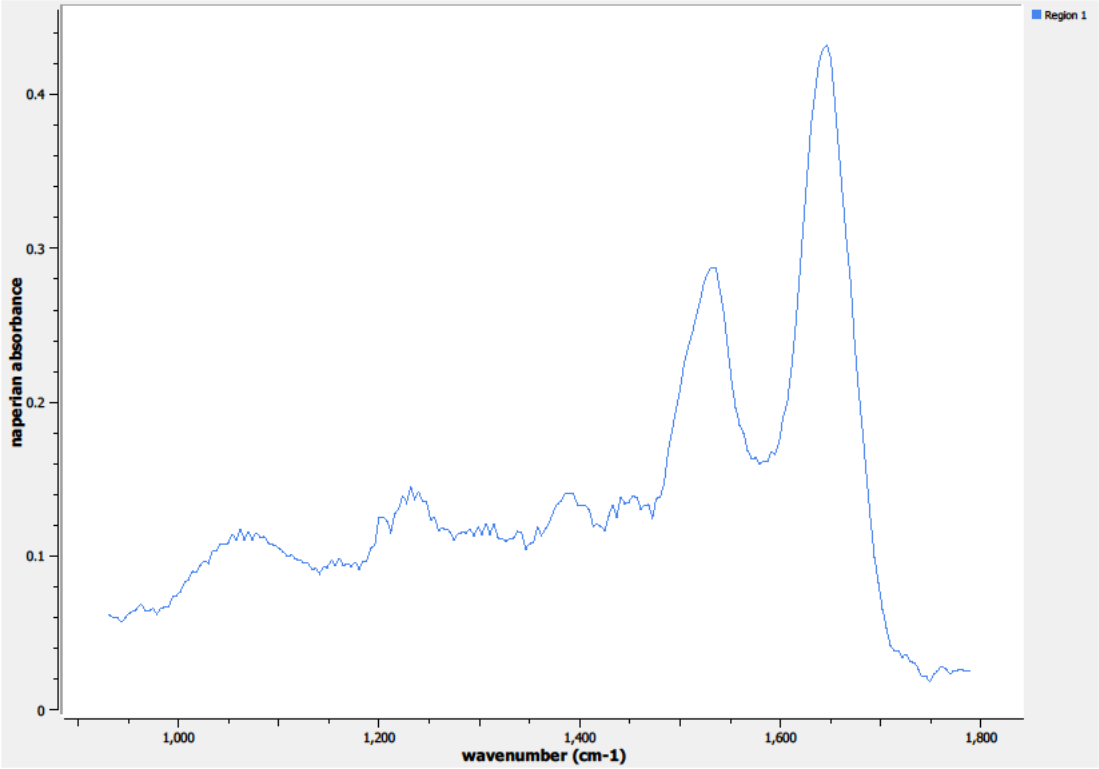



**Figure 0.56. Point D Spectra**

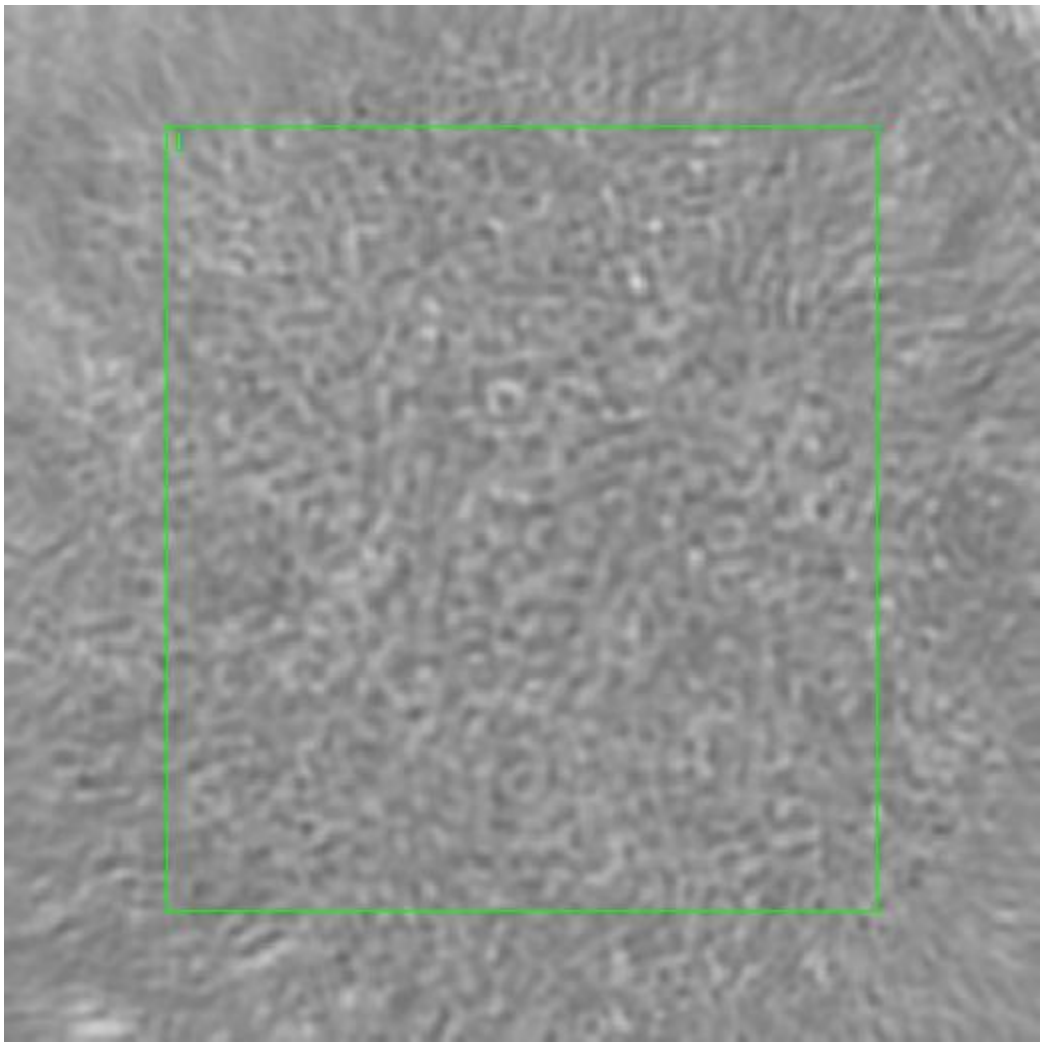


**Figure 0.57. Point E Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

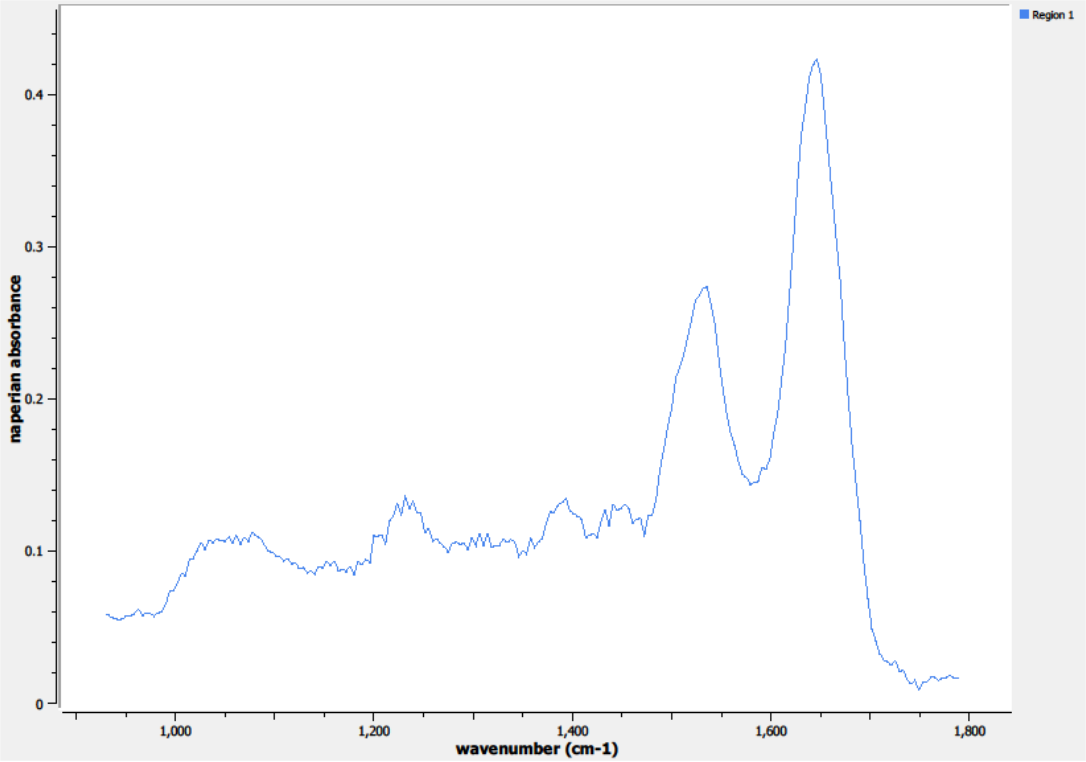



**Figure 0.58. Point E Spectra**

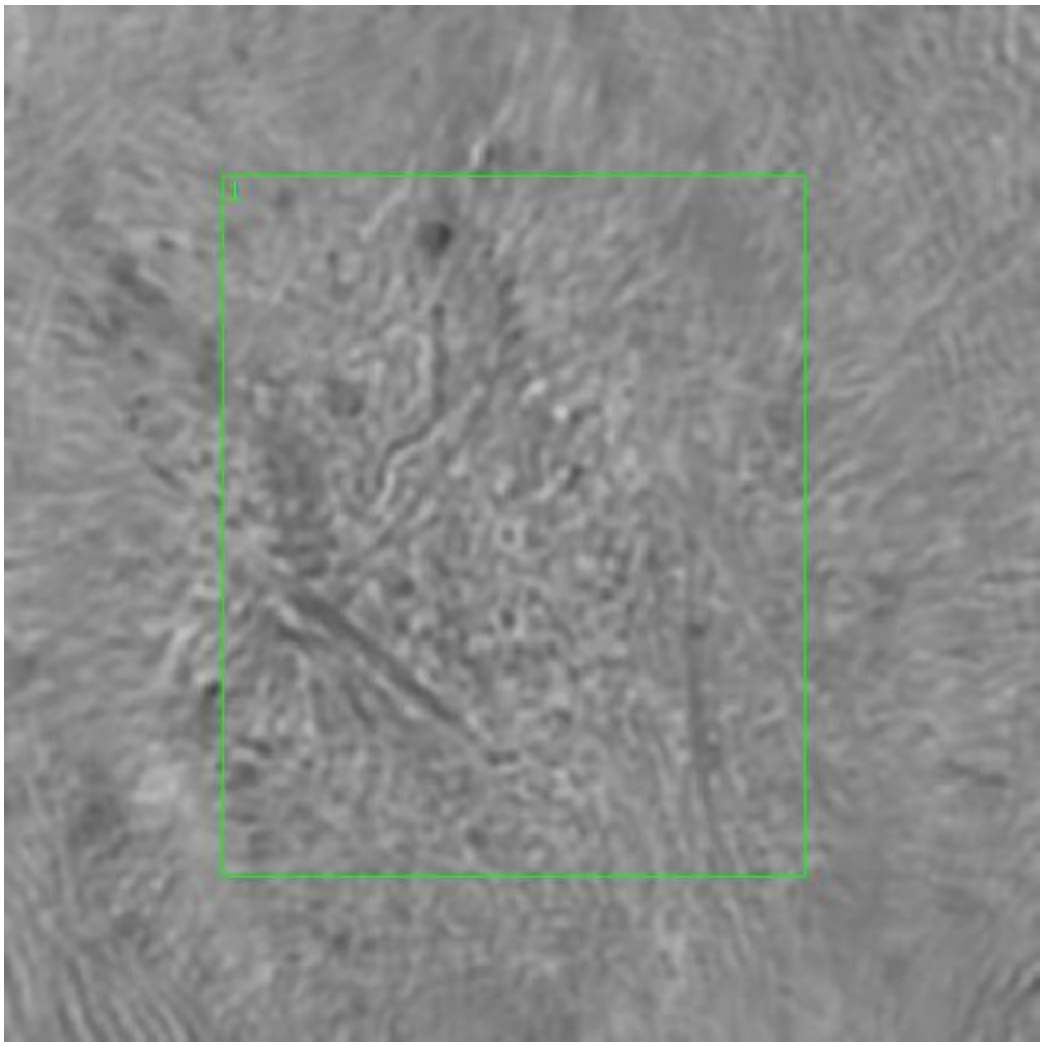


**Figure 0.59. Point F Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

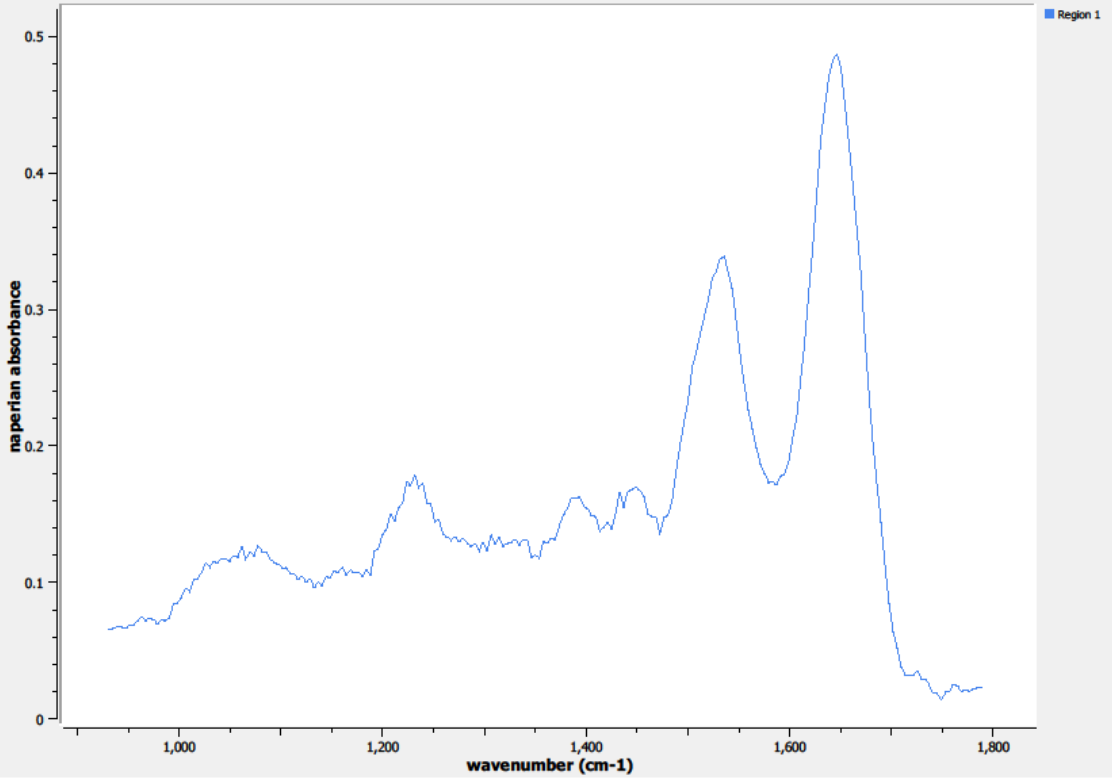



**Figure 0.60. Point F Spectra**

## A8.12-12-A117k

| | |
|---|---|
| Unique ID | 12-12-A117k |
| Age | 67 |
| Race | black |
| Location of Sample Collection | breast, lobular |
| Notes of Interest | malignant lobular adenocarcinoma, invasive lobular carcinoma, ER+, PR+, Her2/neu-, %T (10) / %N (0), unused in data results |

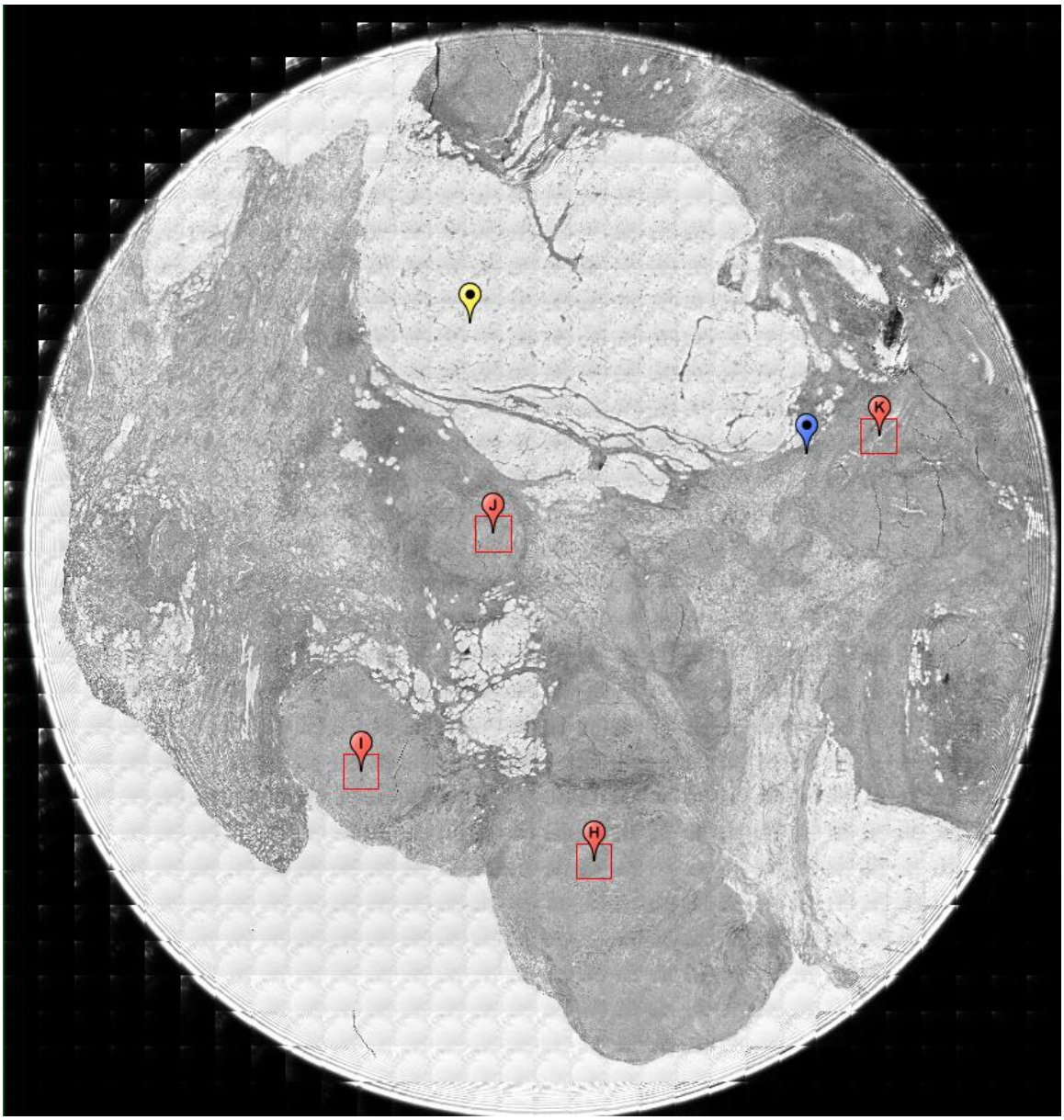



**Figure 0.61. Specimen 12-12-A117k, breast cancer tissue**

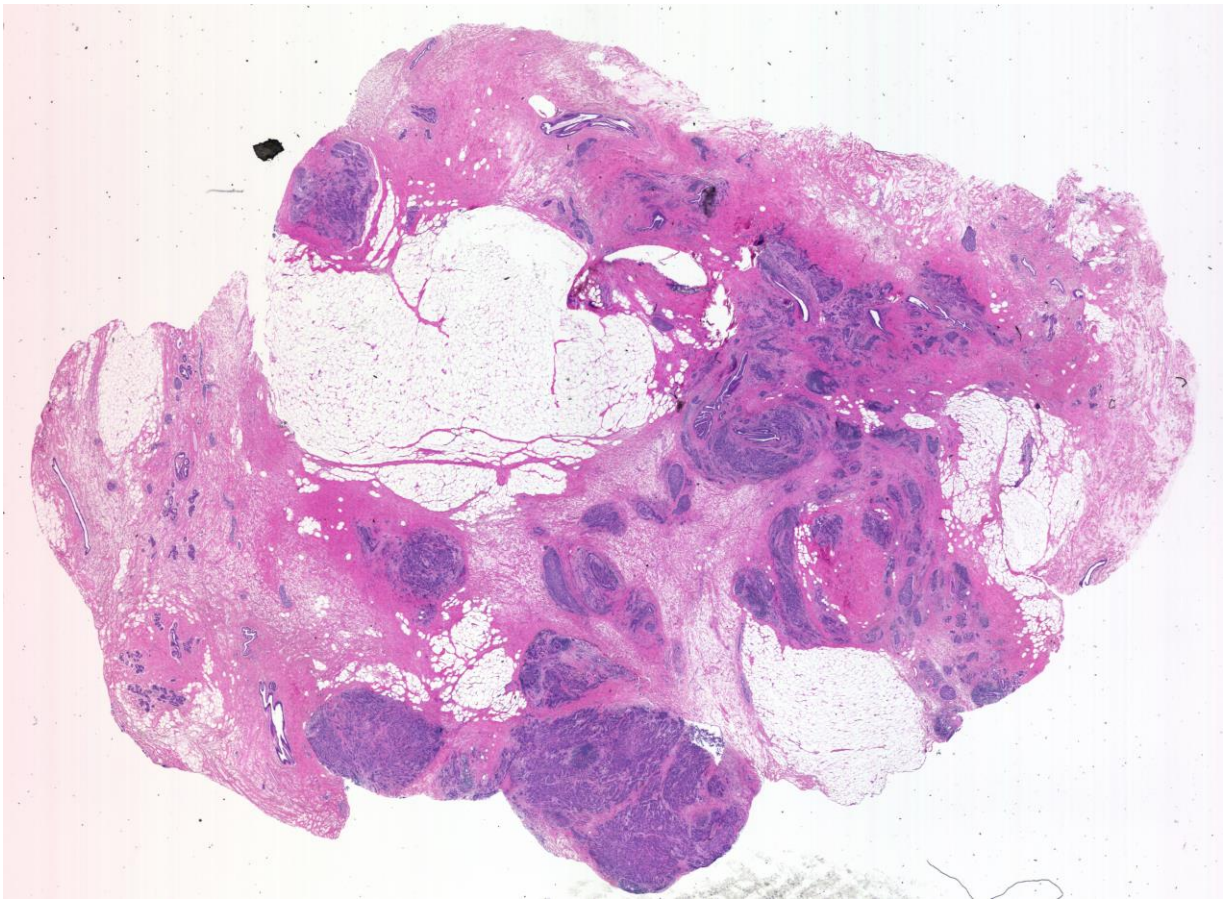


**Figure 0.62. Specimen 12-12-A117k, H&E stain, malignant lobular adenocarcinoma of the breast**

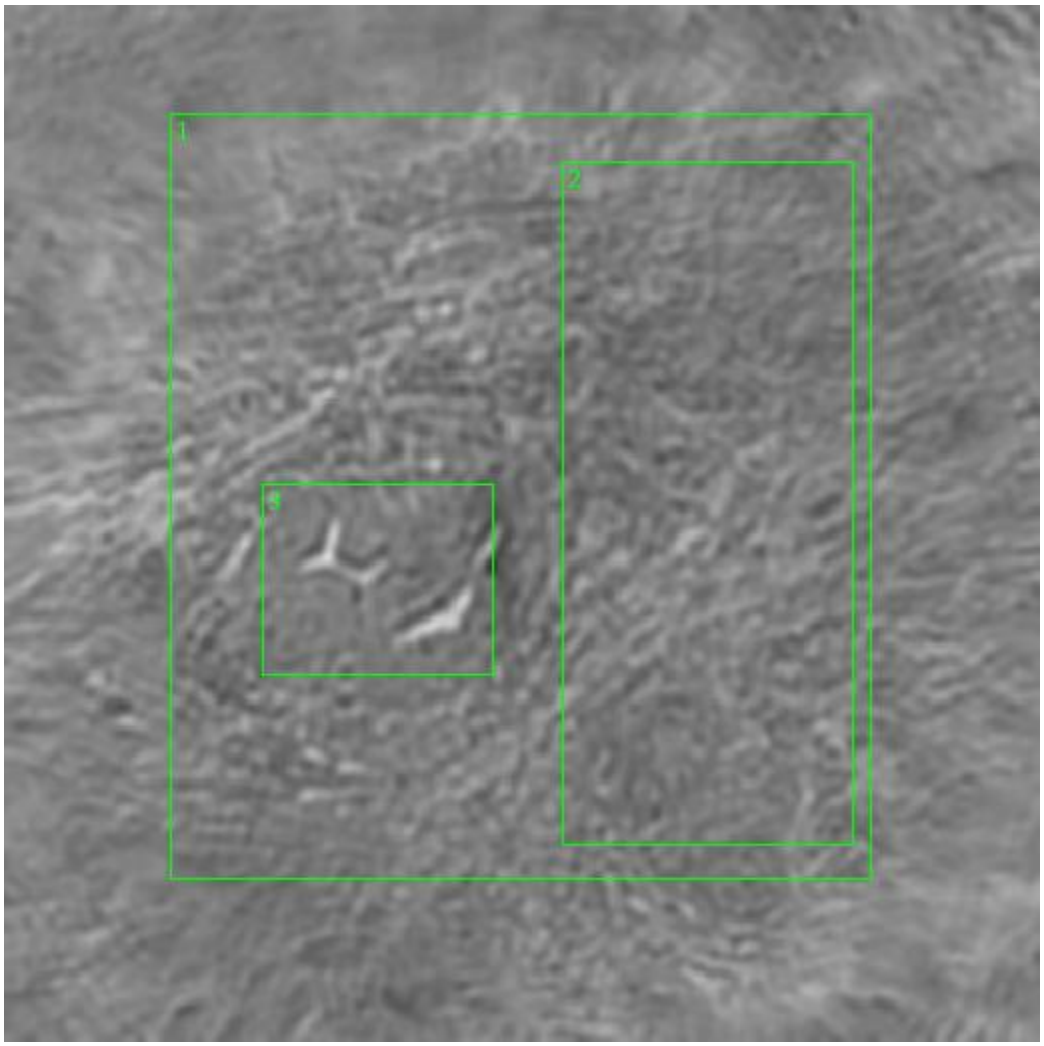



**Figure 0.63. Point H Regions of Interest**
**Regions 1, 2: breast, malignant lobular adenocarcinoma**
**Region 3: breast, malignant lobular adenocarcinoma, concentrated region**

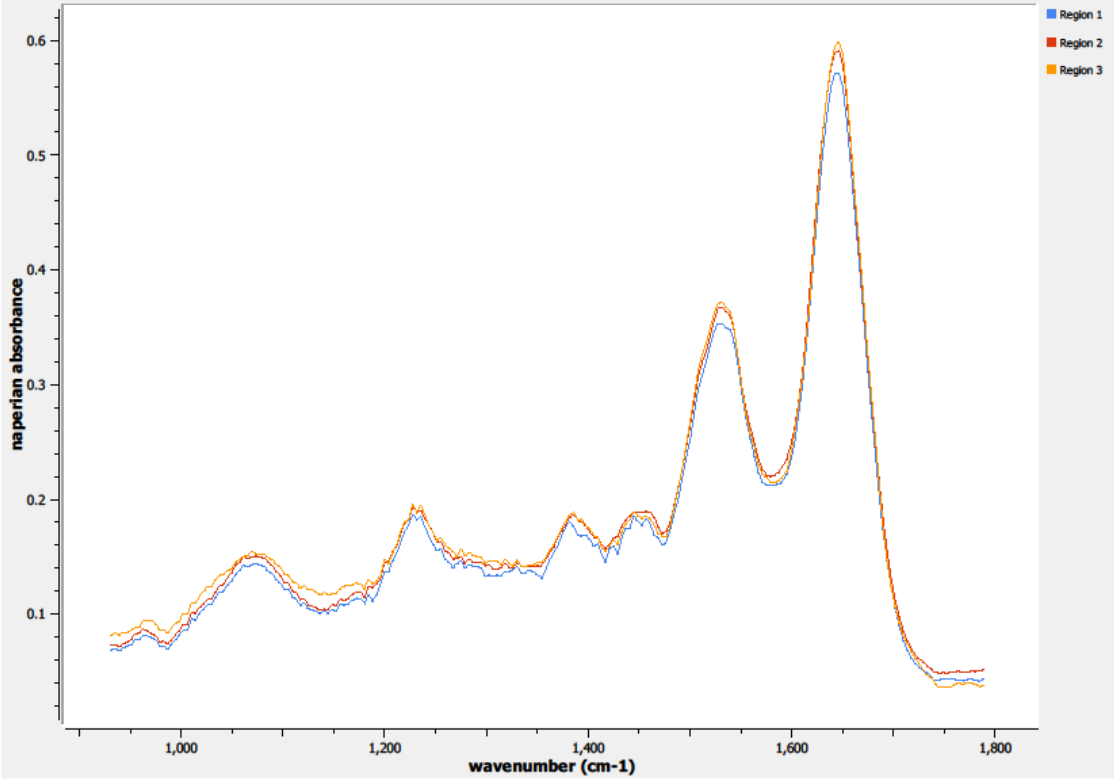



**Figure 0.64. Point H Spectra**

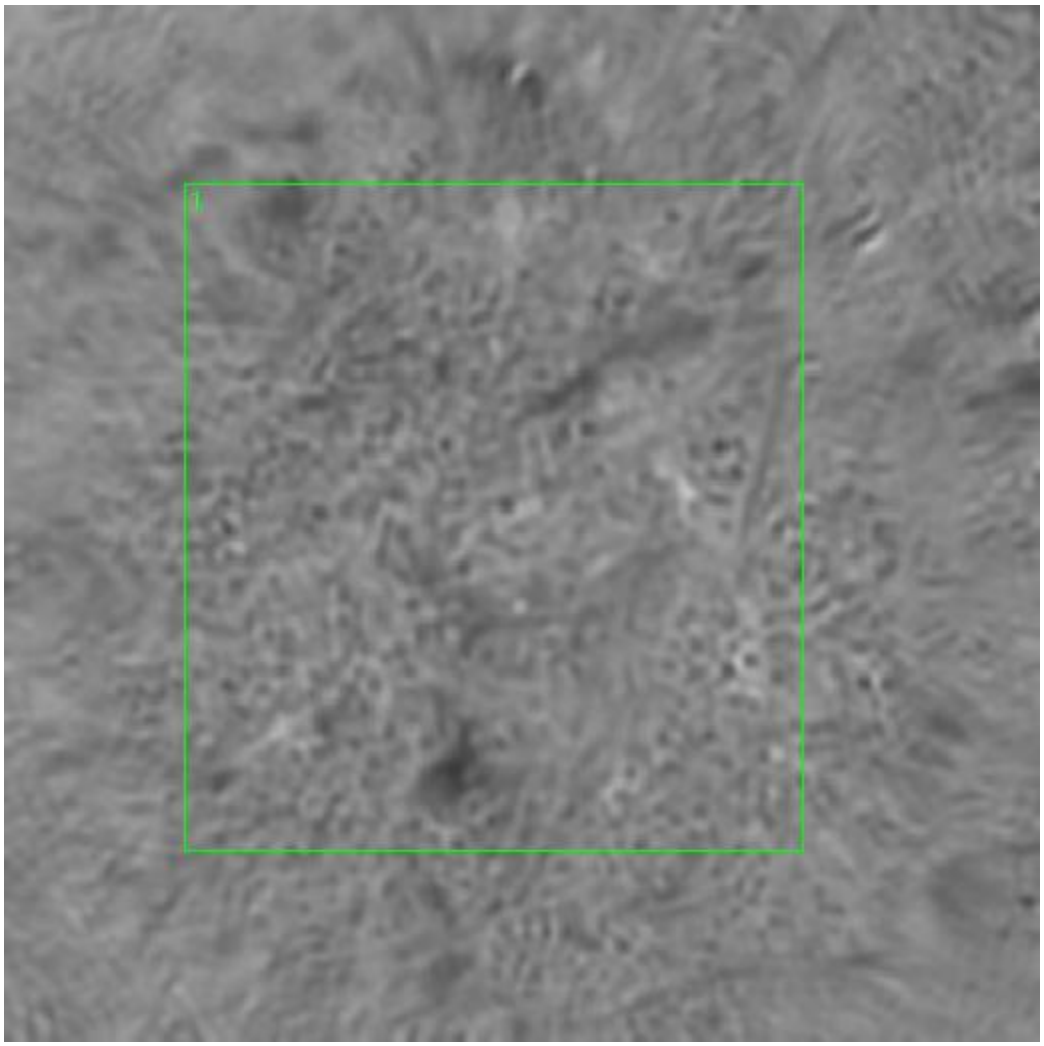


**Figure 0.65. Point I Regions of Interest**
**Region 1: breast, malignant lobular adenocarcinoma**

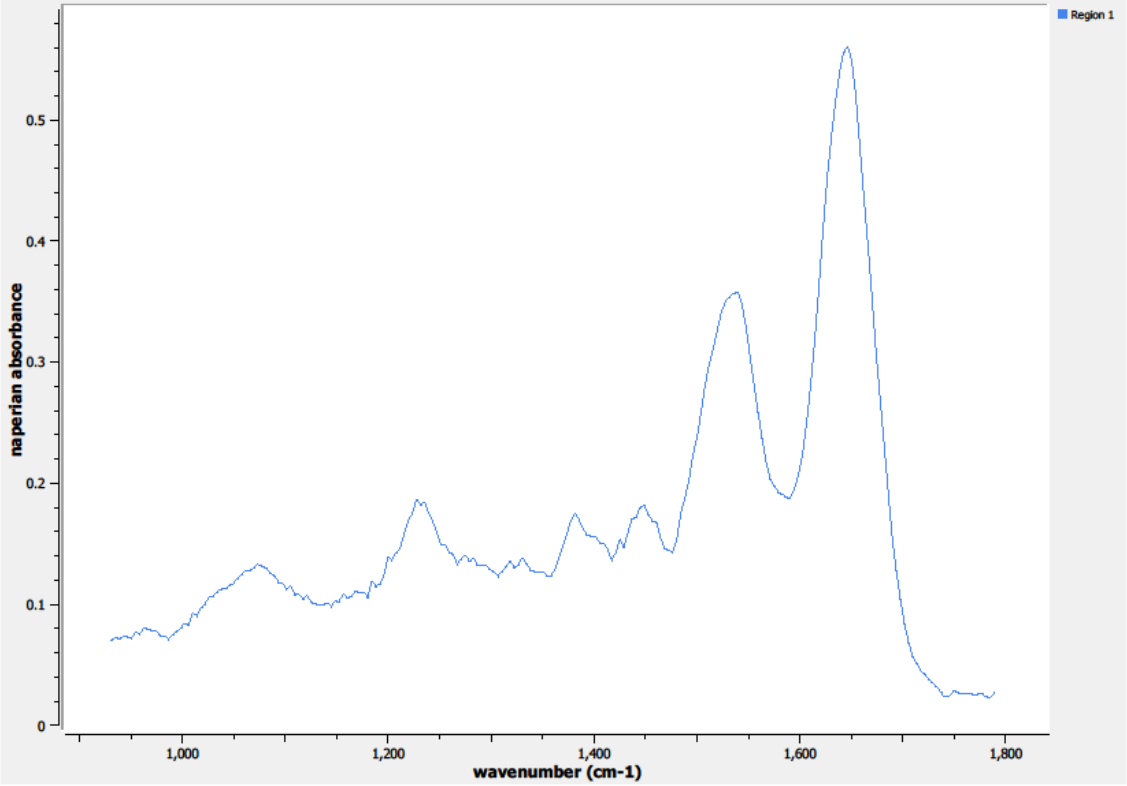



**Figure 0.66. Point I Spectra**

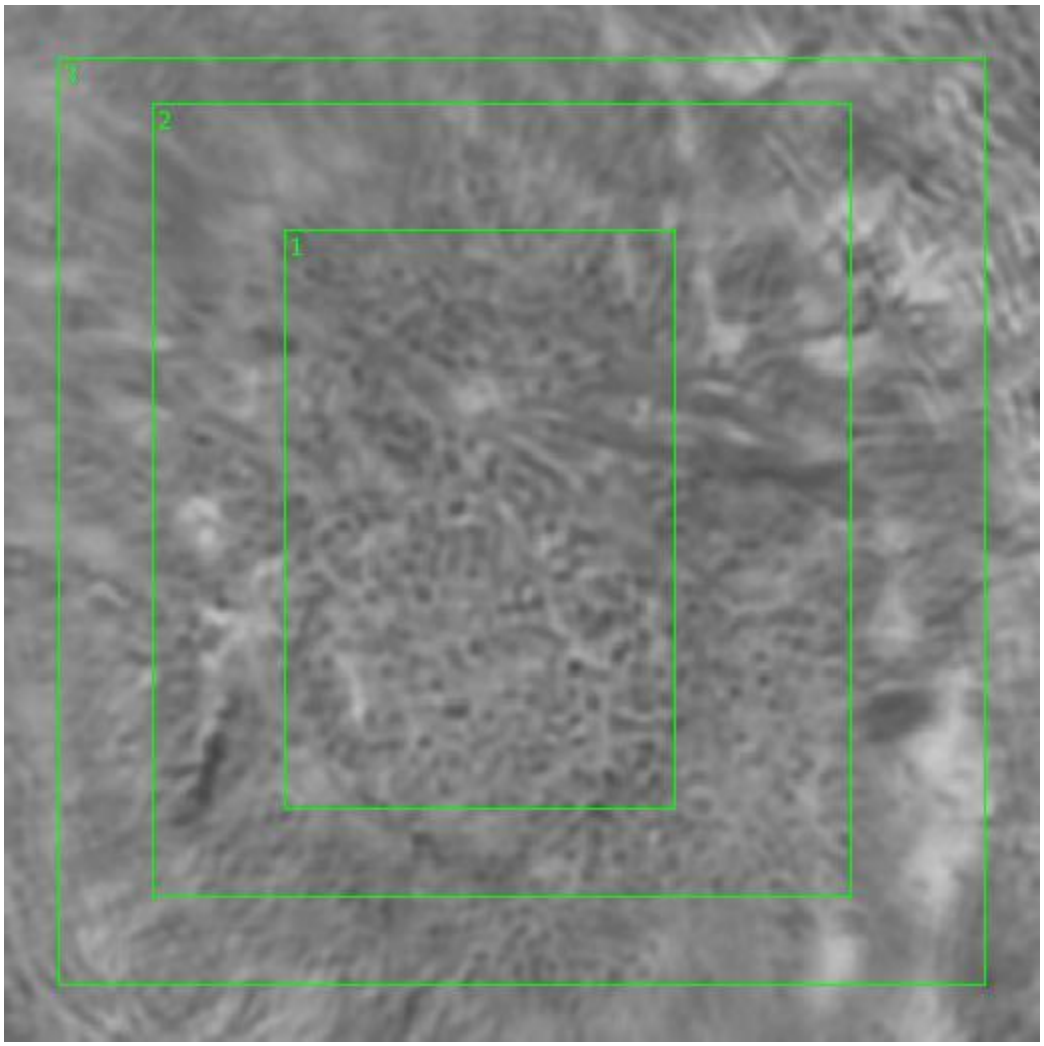



**Figure 0.67. Point J Regions of Interest**
**Regions 1, 2, 3: breast, malignant lobular adenocarcinoma**

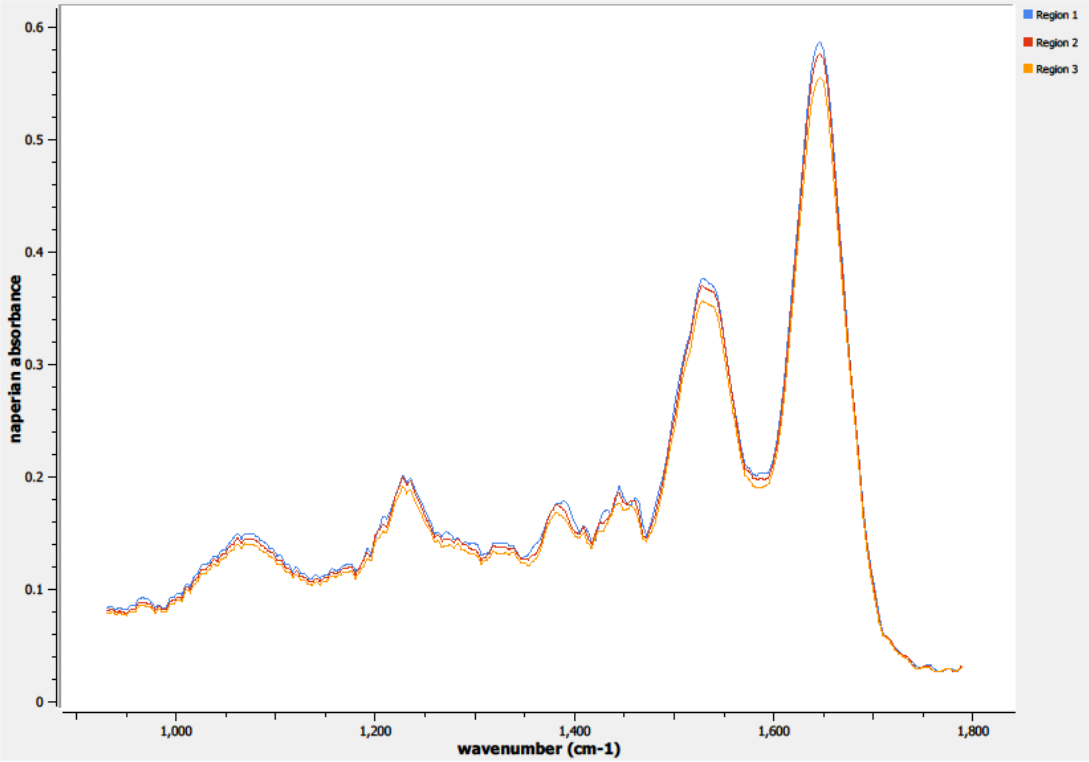



**Figure 0.68. Point J Spectra**

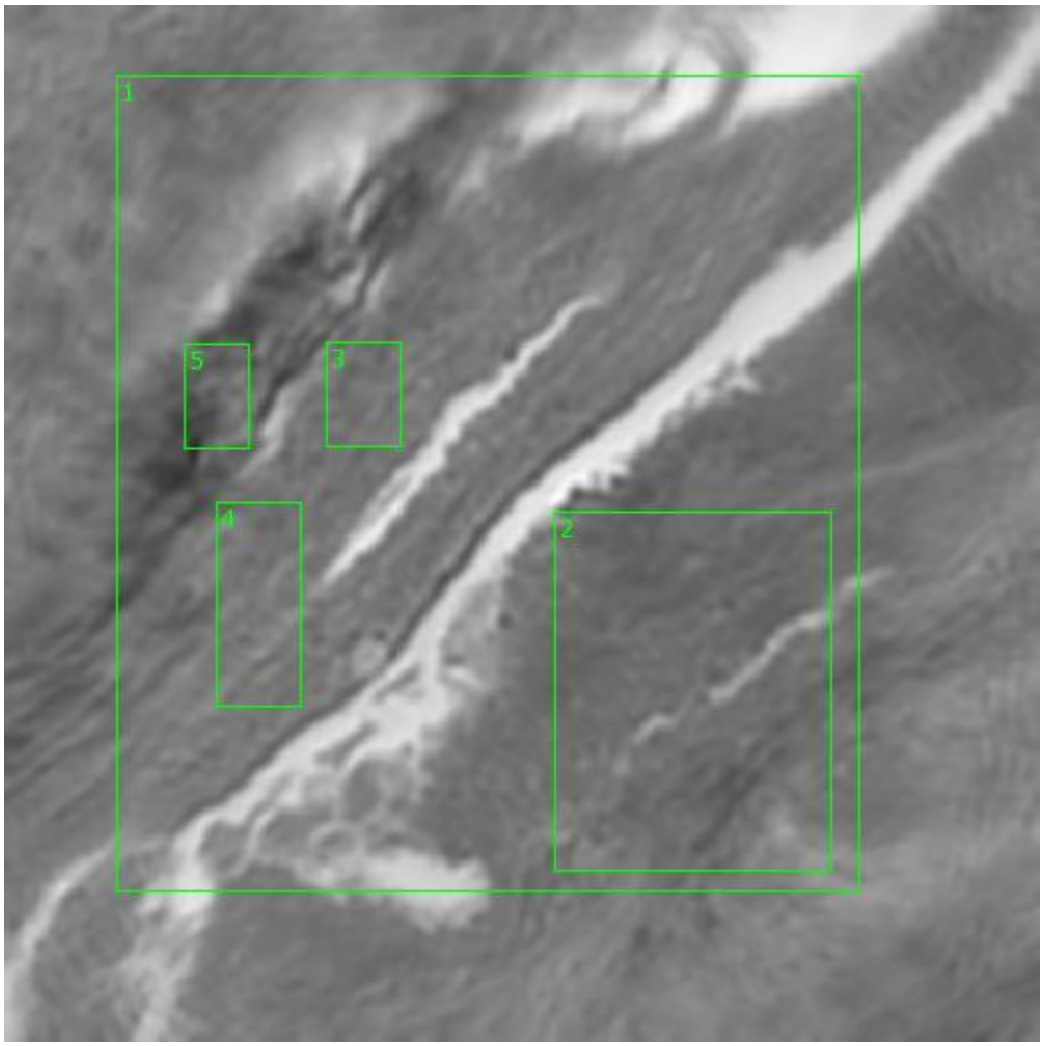



**Figure 0.69. Point K Regions of Interest**
**Regions 1, 2: breast, malignant lobular adenocarcinoma**
**Regions 3, 4, 5: breast, malignant lobular adenocarcinoma, concentrated region**

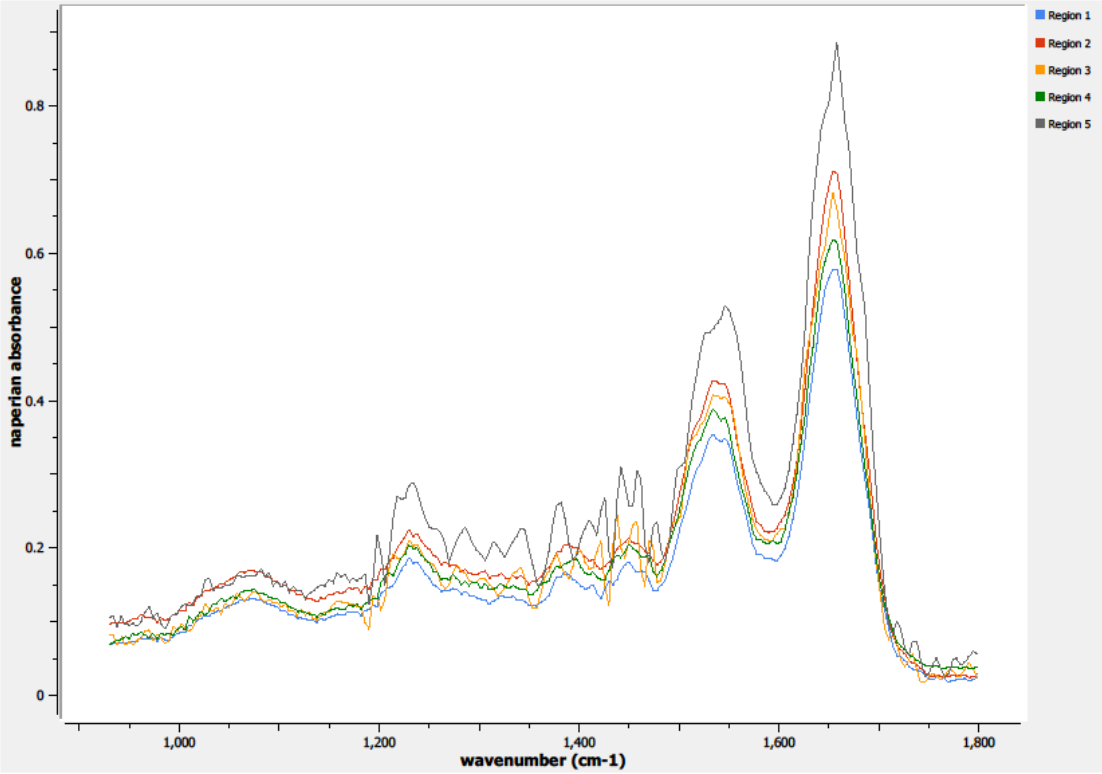



**Figure 0.70. Point K Spectra**

## A9.13-05-A013h

| | |
|---|---|
| Unique ID | 13-05-A013h |
| Age | 64 |
| Race | white |
| Location of Sample Collection | breast, ductal |
| Notes of Interest | malignant ductal adenocarcinoma, invasive ductal carcinoma, ER+, PR-, Her2/neu-, %T (100) / %N (50) |

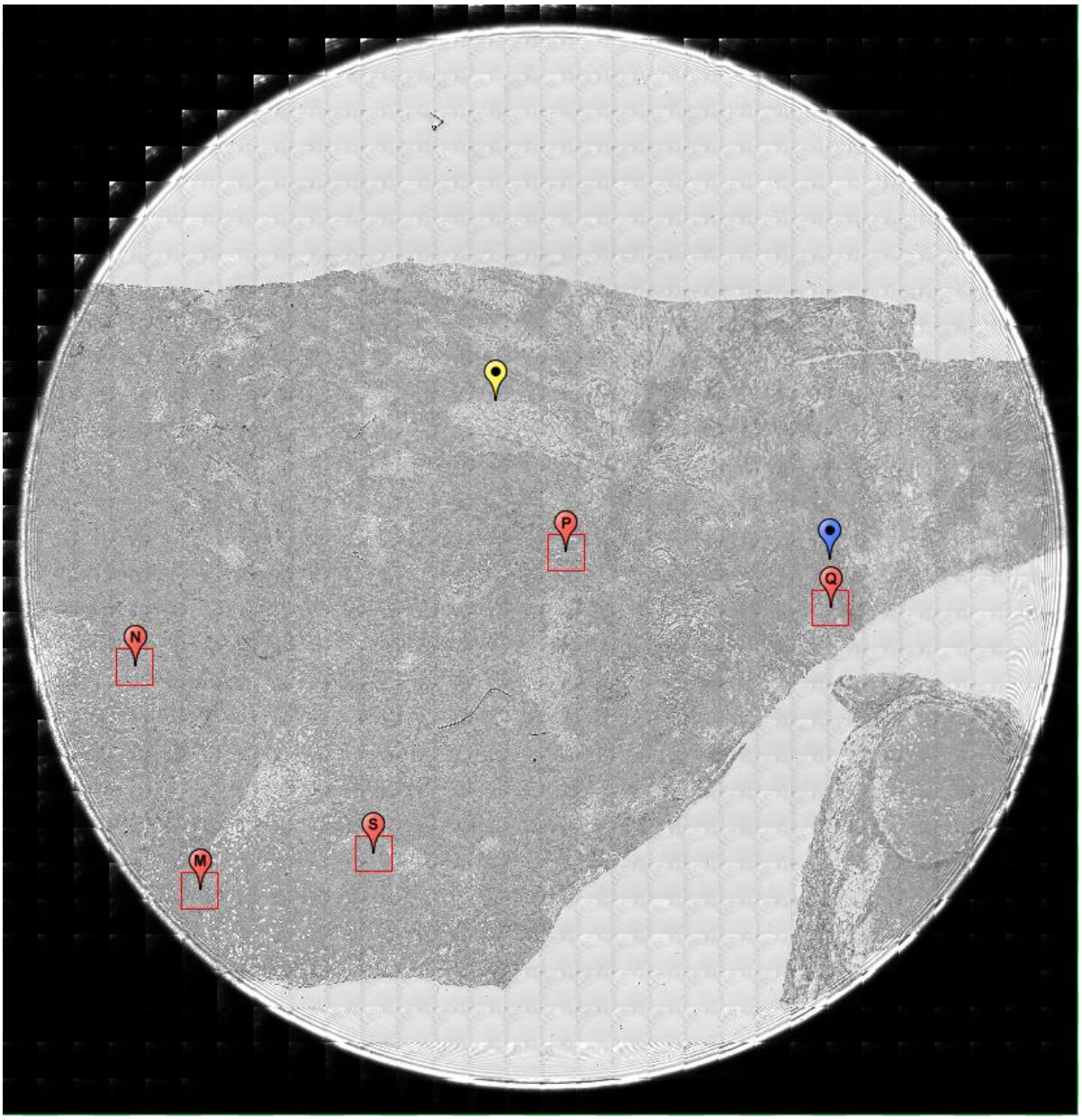



**Figure 0.71. Specimen 13-05-A013h, breast cancer tissue**

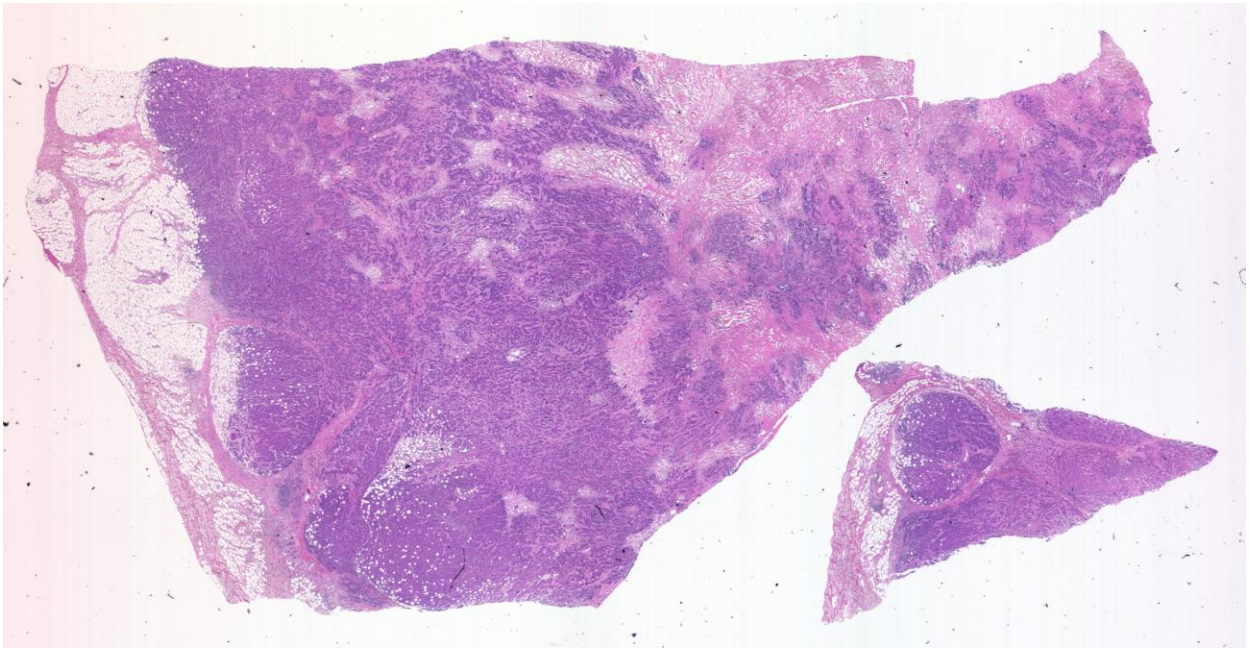


**Figure 0.72. Specimen 13-05-A013h, H&E stain, malignant ductal adenocarcinoma of the breast**

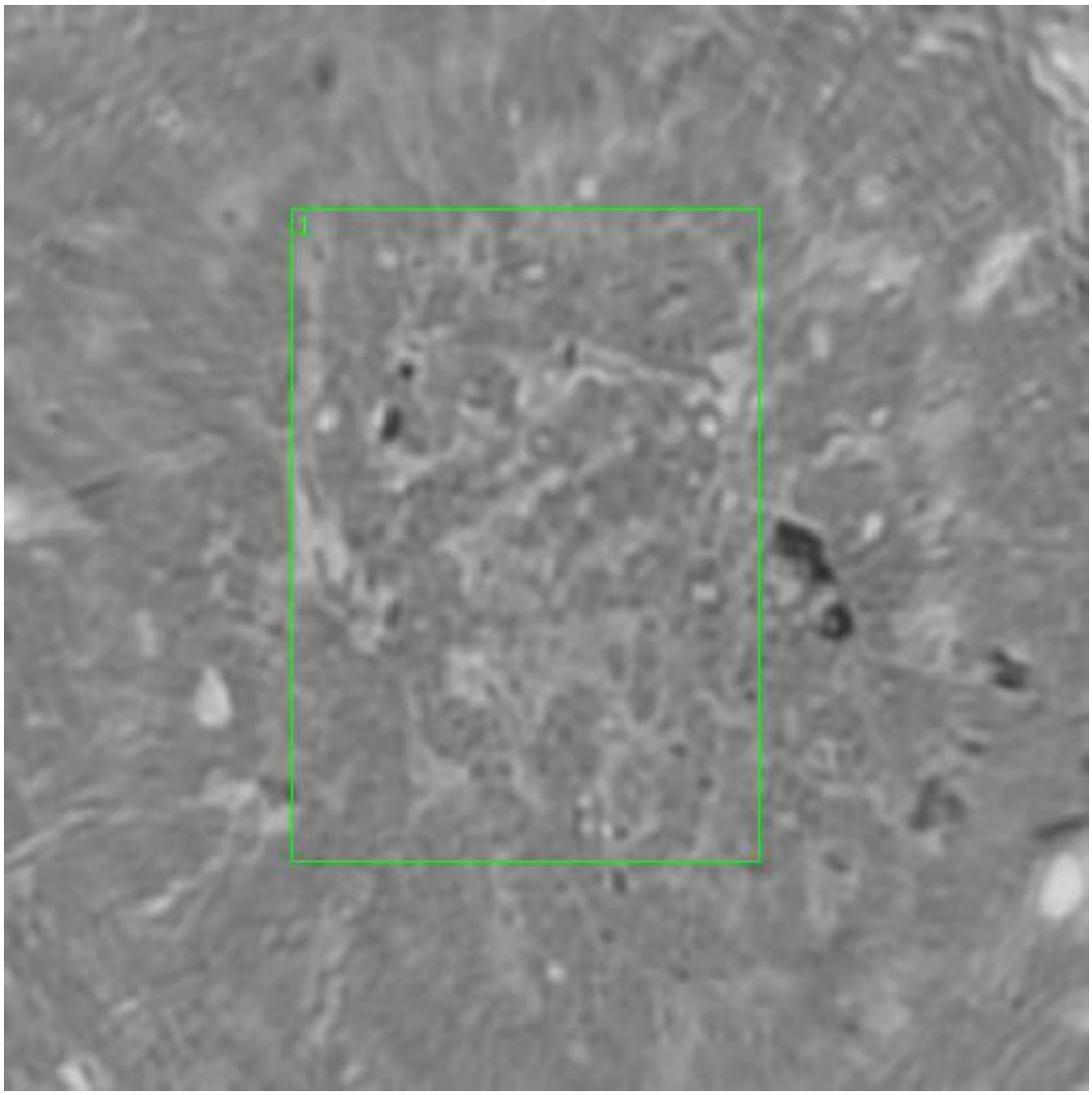


**Figure 0.73. Point M Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

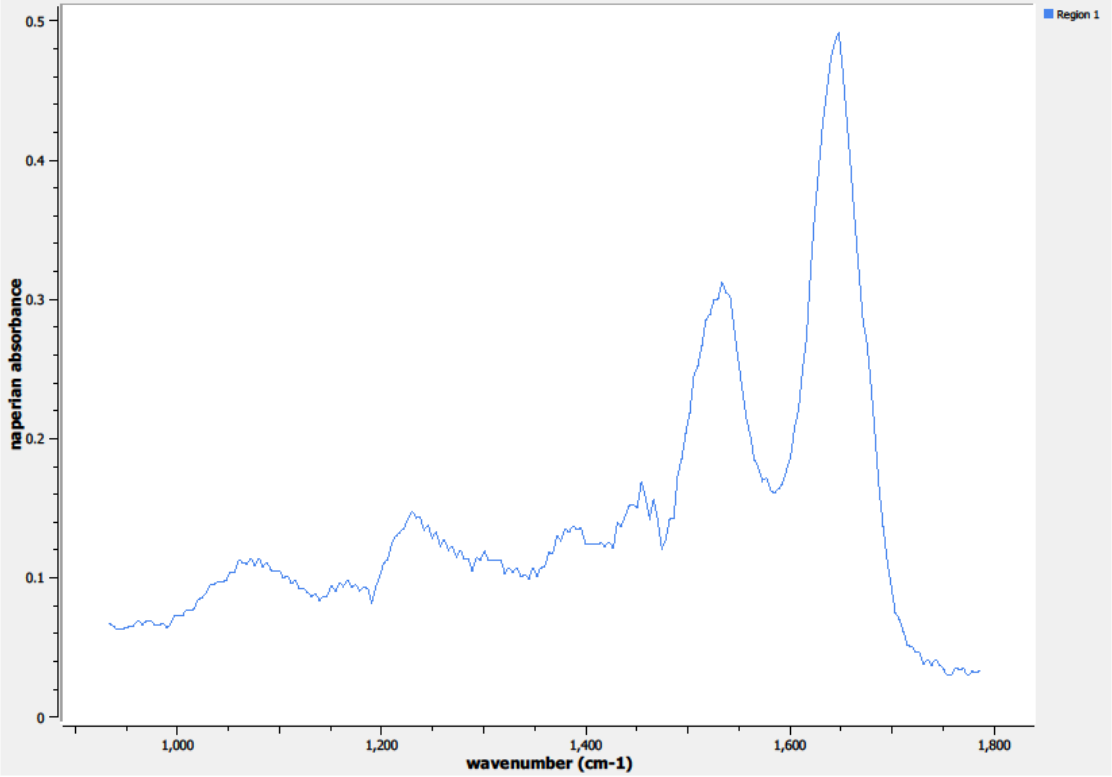



**Figure 0.74. Point M Spectra**

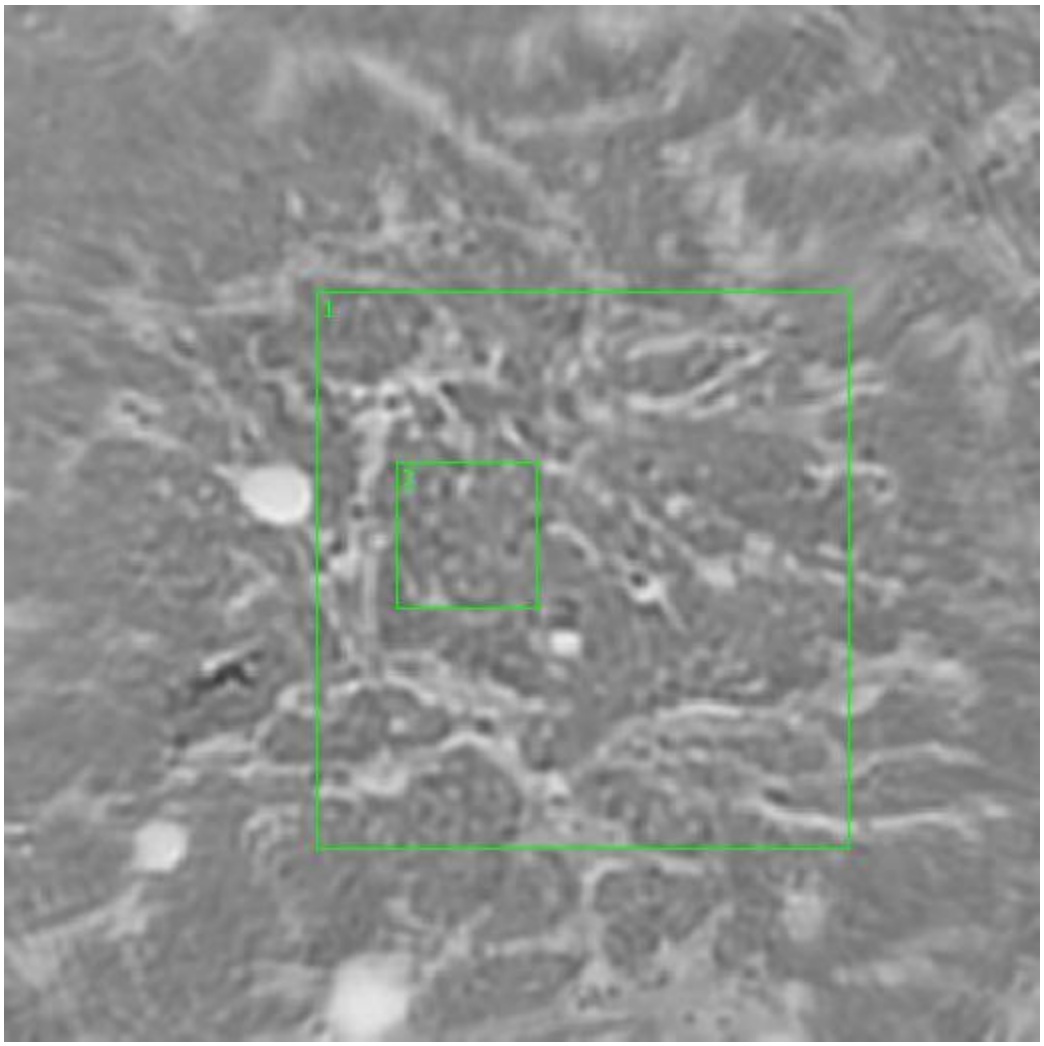


**Figure 0.75. Point N Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**
**Region 2: breast, malignant ductal adenocarcinoma, concentrated region**

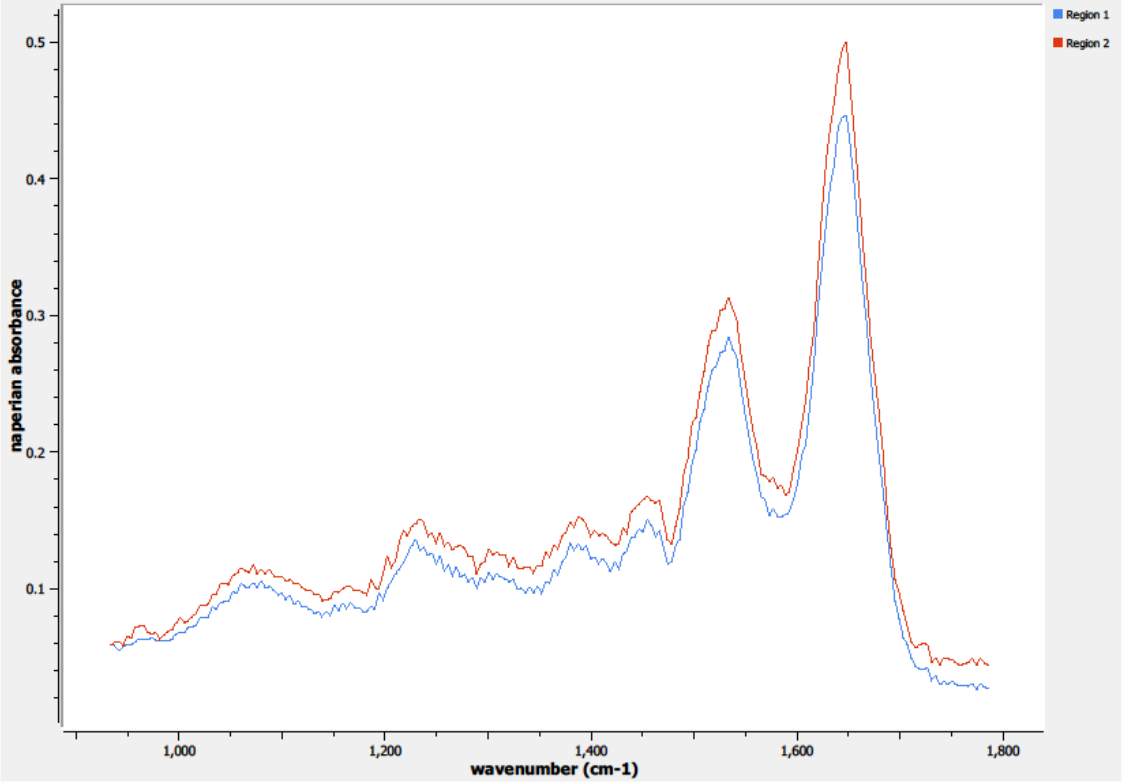



**Figure 0.76. Point N Spectra**

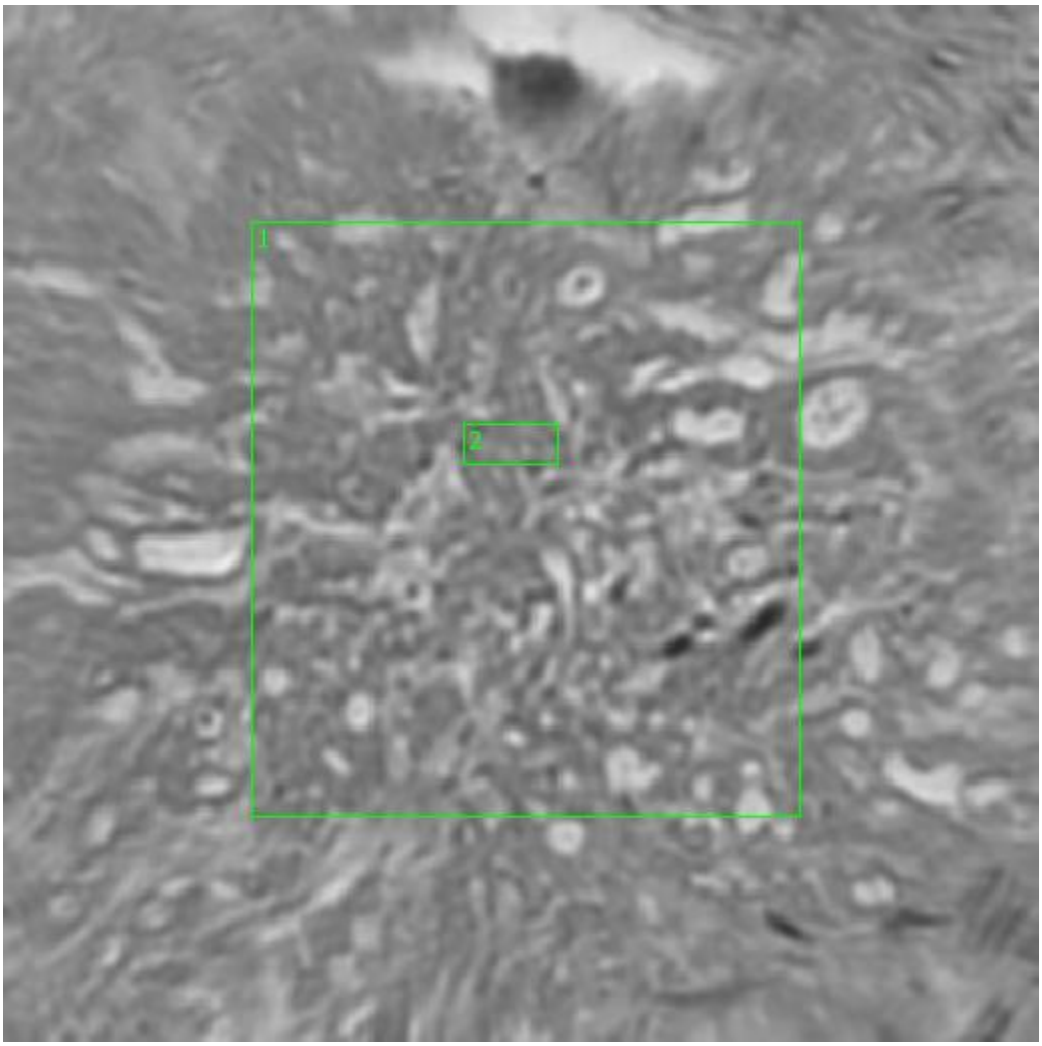



**Figure 0.77. Point P Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**
**Region 2: breast, malignant ductal adenocarcinoma, concentrated region**

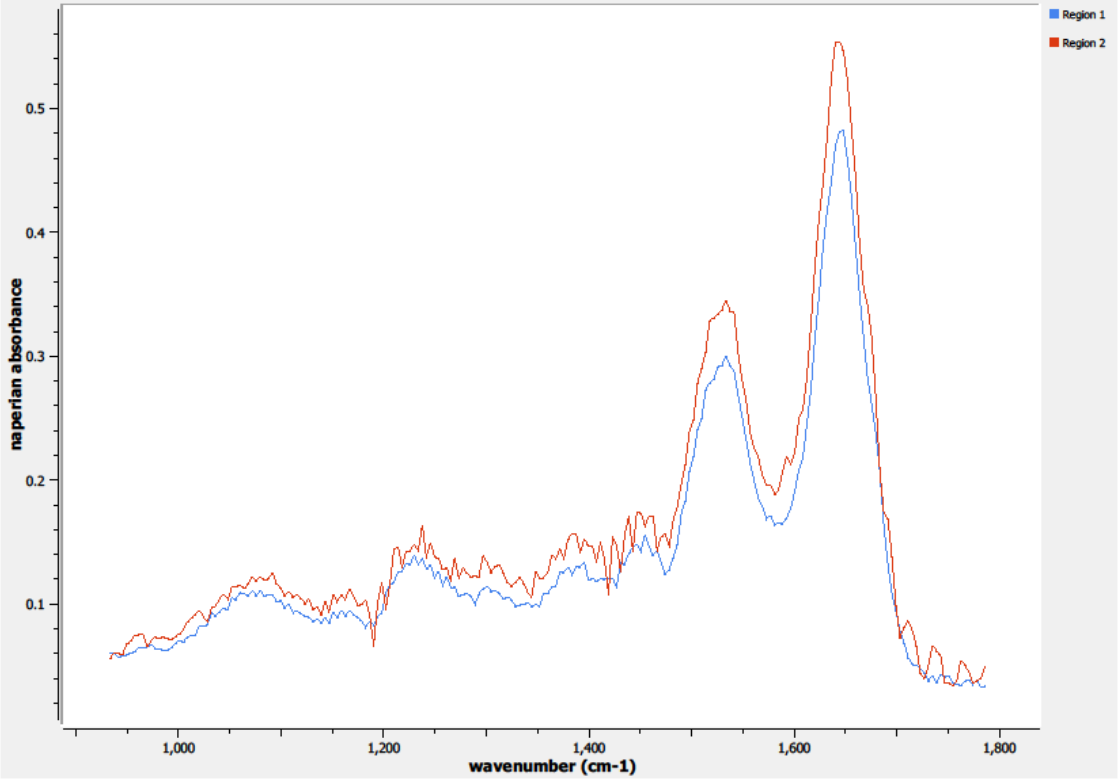



**Figure 0.78. Point P Spectra**

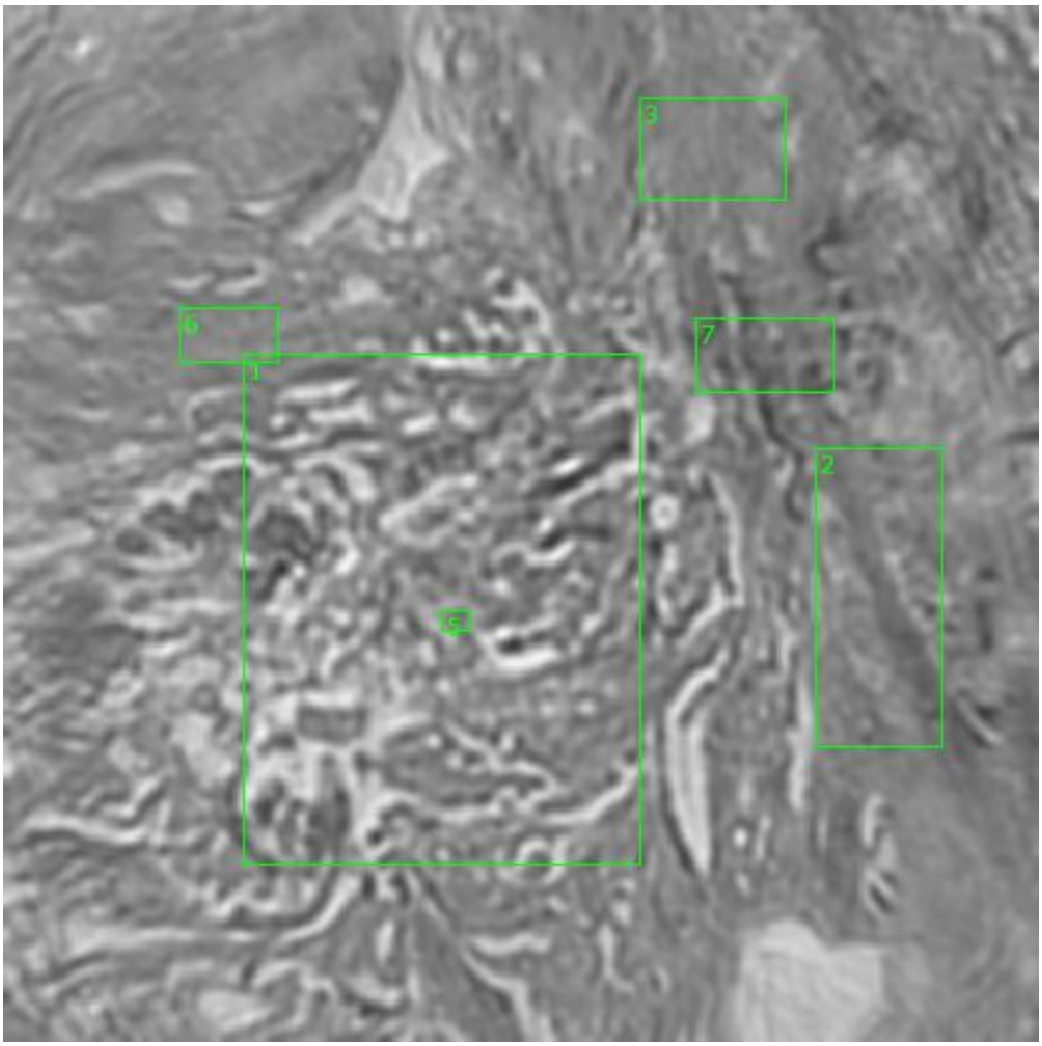



**Figure 0.79. Point Q Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**
**Regions 2, 3, 5, 6, 7: breast, malignant ductal adenocarcinoma, concentrated region**

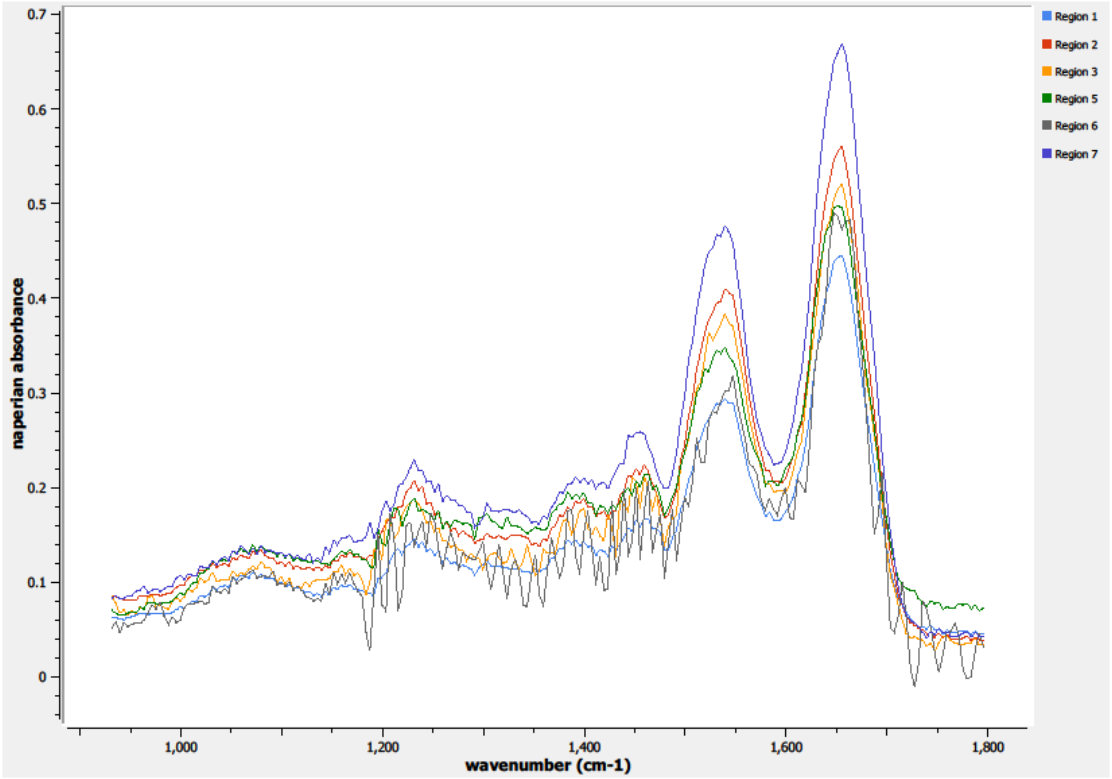



**Figure 0.80. Point Q Spectra**

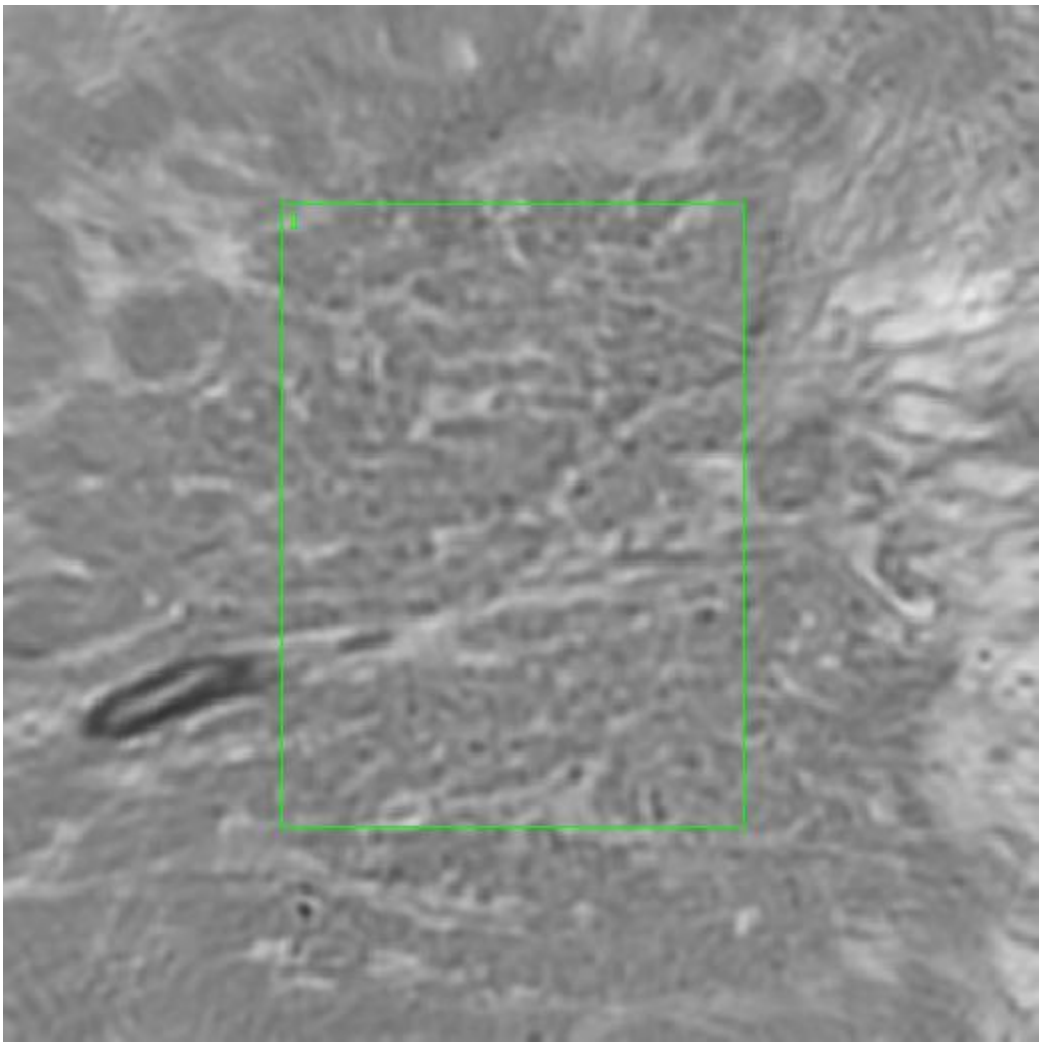


**Figure 0.81. Point S Regions of Interest**
**Region 1: breast, malignant ductal adenocarcinoma**

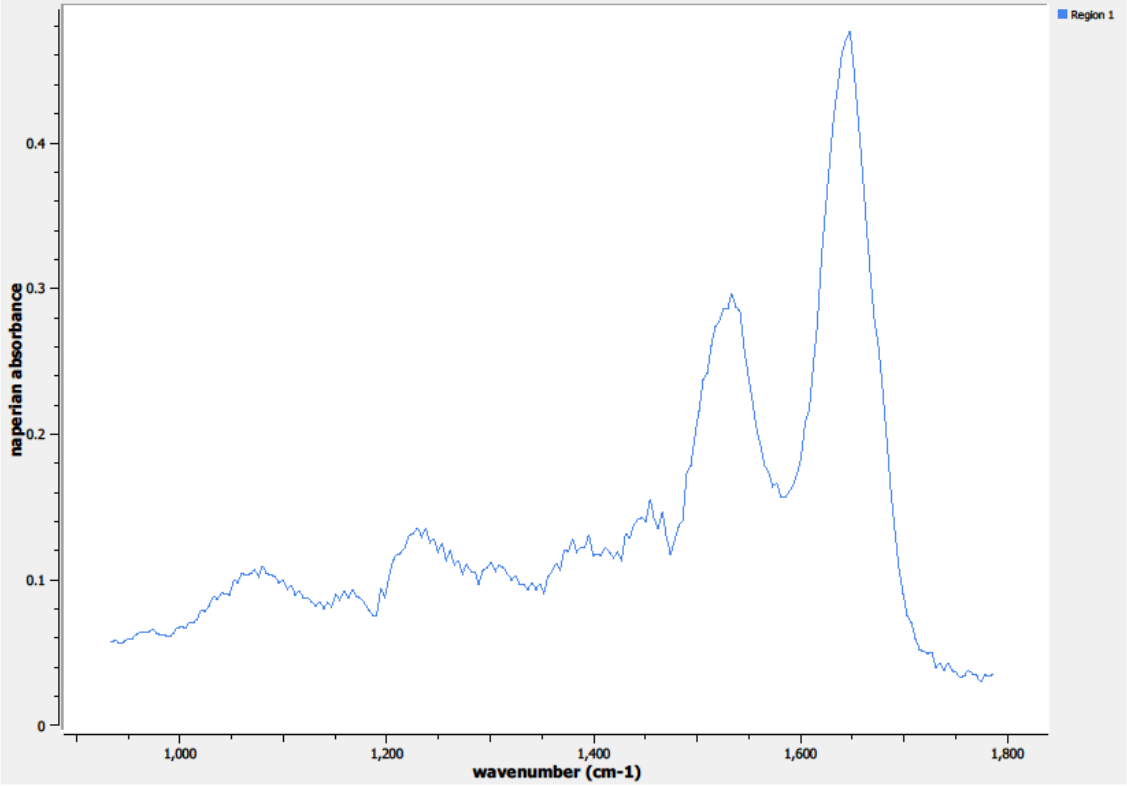



**Figure 0.82. Point S Spectra**